An investigation of the association between physical activity and balance in older adults

Ilona McMullan

MSc Health Psychology

Institute of Nursing and Health Research

Faculty of Life and Health Sciences

Ulster University

Thesis submitted for Doctor of Philosophy (Ph.D.)

December 2018

(I confirm that the word count for this thesis is less than 100,000 words excluding the title page, contents, acknowledgements, summary or abstract, abbreviations, footnotes, diagrams, maps, illustrations, tables, appendices, and references).

This thesis is dedicated to my husband, David, without whose constant and unwavering support I would never have had the courage to begin this journey and see it through to the finish.

“and miles to go before I sleep”

(Robert Frost ‘Stopping by Woods on a Snowy Evening’)

Contents

Acknowledgements

This thesis would not have been possible without the contribution and support from several people. I would like to express my sincere thanks to:

Prof. Suzanne McDonough for her constant support and encouragement throughout the period of research. Also, for her contribution to the independent screening, data extraction, and methodological quality of the studies included in the systematic review (Chapter two).

Prof. Brendan Bunting for his invaluable statistical guidance, input, and patience throughout the whole process. Professor Bunting provided statistical guidance for the analysis of the longitudinal data of TILDA, ELSA, and NICOLA (Chapters three, four, five, and six).

Prof. Mark Tully for helping me recognise my own potential and encouraging me through the process, as well as providing sound advice and guidance. Prof. Tully contributed to the data extraction, and risk assessment for the studies included in the systematic review (Chapter two).

Dr. Karen Casson for her assistance in risk screening, and risk assessment of the studies included in the systematic review (Chapter two). Also, for her 'eagle eye' on editing.

Prof. Margaret Cupples for encouraging me to undertake the journey, and for her contribution to the independent screening of the papers and invaluable feedback on the systematic review (Chapter two).

The Irish Social Science Data Archive for access to the TILDA dataset.

The UK Data Archive for providing data to the ELSA data.

The NICOLA project team for providing access to the NICOLA data.

And finally, to my family, Granny Connor, Mum, Dad, David, Star, and Bel.

Conference presentations and publications

Conference presentations

McMullan II, Bunting BP, McDonough SM, Tully MA & Casson K. (2016). The association between physical activity and balance in an older adult population (≥50 years). Poster presentation at Centre of Excellence for Public Health Northern Ireland Conference, November 2015.

McMullan II, Bunting BP, McDonough SM, Tully MA & Casson K. (2016). Humpty Dumty had a great fall. Poster presentation at Centre of Excellence for Public Health Northern Ireland Conference, November 2016.

McMullan II, Bunting BP, McDonough SM, Tully MA & Casson K. (2017). Factors influencing the mediating effects of physical activity on balance over two years in older community dwelling adults. Oral presentation at the Falls, Frailty and Bone Health Conference, November 2017.

McMullan II, Bunting BP, McDonough SM, Tully MA & Casson K. (2016). Humpty Dumty the real story. Poster presentation at the All Ireland Gerontological Nurse Association, May 2017.

McMullan II, Bunting BP, McDonough SM, Tully MA & Casson K. (2017). Active Times. Poster presentation at Centre of Excellence for Public Health Northern Ireland Conference, November 2018.

Publications

McMullan II, McDonough SM, Tully MA, Cupples M, Casson K & Bunting BP. (2015). Protocol: The association between physical activity and balance in an older adult population (≥50 years) PROSPERO (CRD42016039114).

McMullan II, McDonough SM, Tully MA, Cupples M, Casson K & Bunting BP. (2018). The association between balance and free-living physical activity in an older healthy community-dwelling adult population: A systematic review and meta-analysis. BMC Public Health, 18, 431. Doi: https://doi.org/10.1186/s12889-018-5265-4

Summary

Balance is critical for health and well-being, but the ageing process is a key risk factor for poor balance which may lead to falls, disability, and death in older adults. The benefits of physical activity (PA) are recognised in policy and guidelines for fall prevention, where moderate to vigorous PA (MVPA) is recommended for the general population, and balance and co-ordination exercise is recommended for adults at higher risk of falling (≥65 years). Global adherence to PA guidelines is low, and older adults are more likely to engage in low intensity PA (LPA). However, less is understood regarding its benefits for balance. Furthermore, PA and balance assessment is complex where self-reported measures are subject to bias and guidance on which combination of indirect balance measures is appropriate is lacking.

This thesis aims to contribute to the understanding of the relationship between PA and balance in older adults (≥50 years). A systematic review of the existing literature identified that free-living PA defined as activity for leisure, occupational, and travel was associated with better balance in healthy older adults. A Structural Equation Modelling (SEM) approach using data from the Irish Longitudinal Ageing study (TILDA), and the English Longitudinal Study of Ageing (ELSA) showed that multiple indirect measures provided effective balance assessment; that balance declined by 25%-29% over time; that PA was beneficial to balance where an extra Metabolic Equivalent minute per week improved balance by 5% over two years; that MVPA improved balance in older adults ≤70 years and slowed the rate of decline; and that LPA

improved balance in older adults ≥70 years. An investigation into the robustness and generalisability of the results using the Northern Irish Cohort of Longitudinal Ageing (NICOLA) confirmed that the findings from TILDA or ELSA can be generalised to other studies of ageing.

Abbreviations

ABC	Advanced Balance Confidence scale
ADL	Activities of Daily Living
AIC	Akaike Information Criterion
AMED	Allied and Complimentary Medicine Database
ANCOVA	Analysis of Covariance
BIC	Bayesian Information Criterion
BoS	Base of Support
CAPI	Computer Assisted Personal Interview
CDSR	Cochrane Database of Systematic Reviews
CENTRAL	Central Register of Controlled Trials
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CHAMPS	Communities Health Activities Models Program for Seniors
CI	Confidence Interval
CINAHL	Cumulative Index to Nursing and Allied Health Literature
CMO	Chief Medical Officer
CNS	Central Nervous System
CoM	Centre of Mass
CoP	Centre of Pressure

CSP	**Chartered Society of Physiotherapy**
DALYs	**Daily Adjusted Life Years**
DoH	**Department of Health**
ELSA	**The English Longitudinal Study of Ageing**
Est	**Estimate**
GGAI	**Gateway to Global Ageing Initiative**
HA	**Health Assessment**
HSE	**Health Survey England**
ICC	**Intercorrelation Coefficient**
ICPSR	**Interuniversity Consortium of Political and Social Research**
IPAQ	**International Physical Activity Questionnaire**
ISSDA	**Irish Social Science Data Archive**
LCA	**Latent Class Analysis**
LGM	**Latent Growth Model**
LogMAR	**Minimal Angle of Resolution**
LPA	**Low intensity physical activity**
MANCOVA	**Multivariate Analysis of Covariance**
MD	**Mean Difference**
MEDLINE	**Medical Literature Analysis Retrieval System Online**
MeSH	**Medical Subject Headings**
METS	**Metabolic equivalent**
MLE	**Maximum Likelihood Estimator**

MLR	**Maximum Likelihood Estimation with Robust standard errors**
MLTPAQ	**Minnesota Leisure Time Physical Activity Questionnaire**
MMSE	**Mini Mental State Exam**
MPA	**Moderate intensity physical activity**
MVPA	**Moderate to Vigorous intensity Physical Activity**
NCD	**Non-Communicable Diseases**
NHS	**National Health Service**
NICE	**National Institute for Health and Care Excellence**
NICOLA	**The Northern Irish Cohort of Ageing study**
NOS	**Newcastle Ottawa Scale**
PAGAC	**Physical Activity Guidelines Advisory Committee**
PASS	**Physical Activity Status Score**
PRISMA	**Preferred Reporting Items for Systematic Reviews & Meta-analysis**
PTA	**Pure Tone Audiometry test**
RAPA	**Rapid Assessment of Physical Activity**
RCP	**Royal College of Physicians**
RCTs	**Randomised Controlled Trials**
RMSEA	**Root Mean Square Error of Approximation**
RoI	**Republic of Ireland**
RR	**Relative Risk**
SCQ	**Self-Completed Questionnaire**

SDM	**Standardised Difference in the Mean**
SE	**Standard Error**
SEM	**Structural Equation Model**
SES	**Socio-economic Status**
SMD	**Standardised Mean Difference**
SOT	**Sensory Organisation Test**
SRMR	**Standardised Root Mean Square residual**
TLI	**Tucker Lewis Index**
TILDA	**The Irish Longitudinal Ageing study**
TUG	**Timed Up and Go test**
UK	**United Kingdom**
VPA	**Vigorous intensity physical activity**
WHO	**Worldwide Health Organisation**
YPAS	**Yale Physical Activity Survey**

Abbreviation of initials

IM	**Ilona McMullan**
KC	**Karen Cason**
MC	**Margaret Cupples**
MT	**Mark Tully**
SD	**Suzanne McDonough**

Declaration

Chapter One: Introduction

1.1 Background

Balance, the ability to stay upright and steady whilst moving or stationary, is a complex skill, that requires contribution from neuromuscular, cognitive, and sensory body systems (Horak, 2006; Shumway-Cook & Woollcott, 2001; Sibley et al., 2015; Thomas et al., 2014) (Figure 1.1, page vi). Evidence supports that preserving balance is critical for health and well-being in an ageing population, and that poor balance can lead to accidental falls (Public Health England (PHE), 2018). Accidental falls are defined as the inadvertent activity of coming to rest on the floor or lower level (World Health Organisation (WHO), 2002), and many different biological, environmental, socio-economic, and behavioural risk factors for poor balance have been identified (Bandeen-Roche et al., 2015; Brigola et al., 2015; Bucknix et al., 2015; Chen et al., 2014; 2015; Gillespie et al., 2012; Karlsson et al., 2013; WHO, 2015) (Figure 1.2, page vii).

However, the ageing process has been identified as a key risk factor for falls (Mitchell et al., 2012; WHO, 2007) where, either through disease or degeneration, ageing results in a decline in the systems responsible for balance that commences from the age of 40 years (Chodzko-Zajko et al., 2009). Additionally, where other risk factors are present in an ageing population, then the risk of falls is further increased. For example, several factors have been shown to increase the risk of falls in older adults where older women; those with low socio-economic status (SES); those with poor health or higher rates of comorbid chronic disease or disability; those living in institutional care; and those with a previous history of falling or a fear of falling have a further increased risk of falling (Chen et al., 2015; Cumming et al., 2000; Bandeen-Roche et al., 2015; Bucknix et al., 2015; Karlsson et al., 2013; Stubbs et al., 2015).

Global fall prevalence is high, with 28-35% of those aged ≥65 years, and 32-42% of those aged ≥70 years falling each year, and this presents public health challenges (WHO, 2007; 2015). For example, the individual cost of falls includes injury where 255,000 emergency admissions were reported for those aged ≥65 years, and 173,000 admissions were reported for those aged ≥80 years in the United Kingdom (UK); disability, where falls were reported as the ninth highest cause of disability adjusted life years (DALYs), and the highest cause of injury in the UK; loss of independence, with 20% of UK hip fracture patients entering long-term care in the first year after a fracture; and even death, where UK hip fracture patients post one year have an increased risk of mortality of 18%-33% (Cooper et al., 2010; Forbes et al., 2015; Gillespie et al., 2012; Howe et al., 2011; Karlsson et al. 2013; PHE, 2017; Stubbs et al., 2015; Vieira, Palmer & Chaves, 2016; WHO, 2007).

The economic cost of falls is also high where UK National Health Service (NHS) costs associated with falls are estimated at £2 billion per year (WHO, 2007; 2015). Furthermore, the proportion of people ≥60 years is growing faster than any other age group and is estimated to reach two billion by 2050. As a result, the number of falls is estimated to increase, placing a further burden on already pressured healthcare services for older adults (WHO, 2015).

Global and national physical activity (PA) recommendations and fall prevention policy recognise the role of moderate to vigorous PA for fall prevention. For example, in addition to recommending aerobic activity of moderate intensity (MPA) for at least 150 minutes per week or vigorous intensity (VPA) for at least 75 minutes per week and strengthening exercise on at least two days per week for general health benefits in the general population. Guidance includes balance and co-ordination exercise on at least two days per week for older adults at higher risk of falling such as those aged ≥65 years, with mobility issues, with a history of falls, or with chronic illness (Chartered Society of Physiotherapy (CSP), 2018; Department of Health (DoH), 2011; National Institute for Health and Care Excellence (NICE), 2013; PHE, 2017, 2018; Royal College of Physicians (RCP), 2015; WHO, 2007, 2010, 2015, 2018).

Guidance is less explicit for balance exercise for older adults at lower risk of falling (CSP, 2018; DoH, 2011; NICE, 2013; PHE, 2017; WHO, 2007; 2010). Additionally, guidance focuses on moderate to vigorous intensity activity (MVPA), but globally, adherence to PA guidelines is low (Bann et al., 2015; Milton et al., 2018; Murtagh et al. 2015; Schutzer & Grave, 2004). Moreover, evidence suggests that older adults are more likely to engage in low intensity PA (LPA), that is not exercise, due to health-related issues or physical capability (Ainsworth et al., 2011; Arnardottir et al., 2013; Franco et al., 2015). However, robust evidence supporting the benefits of LPA such as walking for leisure, light housework, or light gardening activities for balance performance is lacking

(Bauman et al., 2016; NICE, 2013; PAGAC, 2018; Schmidt et al., 2017).

In summary, fall prevention is an important health concern for an ageing population. Whilst global and national guidelines recognise the importance of MVPA for older adults at higher risk of falling, guidance is lacking for older adults at lower risk of falling. Furthermore, older adults are more likely to engage in PA that is not MVPA. Consequently, opportunities for fall prevention are potentially being missed (Gillespie et al., 2012; Lamb et al., 2011; RCP, 2015; Sherrington et al., 2017). Therefore, a better understanding of the effects of PA that is not just MVPA is needed (Cooper et al., 2010; 2011; The Physical Activity Guidelines Advisory Committee, 2018).

The following sections provide a brief overview of the current research relating to PA and balance, emphasising the gaps and need for further research.

1.2 Physical activity (PA) overview

1.2.1 Defining physical activity (PA)

PA, movement produced by the skeletal muscles of the body resulting in the expenditure of energy above the basal metabolic rate (Chodzko-Zajko et al., 2009), has multiple ways in which it can be characterised. For example, it can be characterised by the domain in which the activity takes place (leisure, travel, and occupational) (Figure 1.3, page viii). Leisure activity is based on personal interests and needs such as walking, hiking, gardening, swimming, sport, and dance, or planned exercise in the context of daily, family, and community activities (e.g. walking programmes, swimming clubs, or Tai Chi clubs); travel activity includes activities such as cycling or walking; and occupational activity includes activities such as labouring, gardening, and heavy lifting (Howley, 2001; WHO, 2011; 2013).

PA can also be characterised by activity type either by a specific activity such as Tai Chi or within the context of physiological demands such as aerobic, strength, balance and co-ordination exercise activity. However, there is a lack of consensus as to which types of PA contribute to the classifications of aerobic, strength, balance and co-ordination training. For example, walking can be classed as both aerobic and balance training, and so there is an overlap across classifications (Milton et al., 2018).

PA can also be characterised by frequency such as the number of sessions per period; by duration such as the time spent on the activity per session; and by intensity, the measured amount of energy expended carrying out the activity. Intensity can be expressed by, for example, PA type and duration such as steps per minute, or more commonly expressed in Metabolic Equivalents (METs), the ratio of an individual's working metabolic rate relative to their resting metabolic rate (classified as low/light intensity (LPA) (1.6-2.9 METS) (e.g. walking for leisure, light housework, light gardening), moderate intensity (MPA) (3-6 METS) (e.g. brisk walking, dancing, gardening, sports), and vigorous intensity (VPA) (≥ 6 METS) (e.g. running, fast cycling, aerobics, competitive sport) (Ainsworth et al., 2011; WHO, 2018).

1.2.2 Adherence to UK physical activity (PA) guidelines in older adults

UK PA guidance for older adults recommends 150 minutes per week of aerobic activity of MPA (e.g. cycling, brisk walking), or 75 minutes of VPA (e.g. running, swimming, football) as well as muscle strengthening activity on at least two days per week. Additionally, guidance suggests that older adults who are at higher risk of falling (≥ 65 years) should also engage in balance and co-ordination exercise for fall prevention on at least two days per week (CMO, 2011; DoH, 2011).

Despite the benefits of PA to health, global adherence to the guidelines is low (Milton et al., 2018; Murtagh et al., 2015, PAGAC, 2018). For example, in the UK, only 31% of adults aged 19-64 years, and 12% aged ≥65 years met both aerobic and muscle-strengthening guidelines (HSE, 2017). Barriers to adherence include lack of awareness or belief in the benefits of PA; fear regarding personal security; lack of time; lack of social support; lack of interest; as well as environmental issues such as the weather or lack of appropriate facilities (Baert et al., 2011; Cavill & Foster, 2018; Chao et al., 2000; Franco et al., 2015; Schutzer & Graves, 2004). Evidence also suggests that up to 60% of older adults are physically incapable of achieving the guidelines because the guidelines are too physically demanding (Franco et al., 2015; Schutzer and Graves, 2004), with many older adults more likely to engage in LPA (Bauman et al., 2010).

For example, Arnardottir et al. (2013) (n=579 participants), using self-reported measures, suggest that older adults spend 74.5% of non-sleeping time carrying out sedentary behaviour, 21.3% carrying out LPA (defined as 1.8-2.9 METS), and less than 1% carrying out MPV or VPV (>3.6 METS). Additionally, Varma et al. (2014) (n=187 participants), using an objective measure of steps, found that 97.9% of total daily activity of older adults was LPA (>0 steps/minute to <100 steps/min).

Evidence on the benefits of LPA associated with general health are starting to emerge. Tse et al.'s (2015) systematic review of older adults ≥65 years (n=15 studies) found that LPA improved both physical and cognitive health. Additionally, Demakakos et al.'s (2010) analysis of longitudinal data (n=7,466 participants) found that LPA reduced the risk of type two diabetes for adults ≥70 years (hazard ration (HR) 0.53, 95% CI 0.28, 1.02, p=0.059). In addition, Dohrn et al.'s (2018) Swedish cohort study with a 15-year follow-up (n=851 participants) using an objective measure of PA, found that LPA reduced the risk of all-cause mortality by 11% and cardiovascular mortality by 24% in older adults with an average age of 66.7 years (SD 10.2). However, whilst the understanding of the health benefits of LPA are emerging, robust evidence supporting the effects of LPA on balance is needed. For example, cross-sectional studies suggest that objectively measured LPA such as walking for leisure, light housework, or light gardening may prevent falls in higher risk of falling older adults where LPA (>0 step/min to <100 step/min) improved self-reported mobility in older females (≥65 years) with a 14% reduction in the odds of reporting difficulty walking one mile (OR=0.86, 95% CI 0.77, 0.97), and a 16% reduction in the odds of reporting difficulty with one flight of steps (OR=0.84; 95% CI 0.73, 0.98) (n=187 participants; Varma et al., 2014). LPA (1.8-2.9 METS) was also found to have a positive effect on lower extremity measures in adults (≥75 years) with medical conditions (n=802 participants; Osuka et al., 2015). Additionally, LPA (2.0-2.9 METS) was found to improve measures of functionality in older females (n= 290 participants; Izawa et al., 2017). Although these studies can provide an indication of the correlation between PA and indirect measures of balance, they cannot infer causality (PAGAC, 2018; Schmidt et al., 2017).

In summary, whilst PA guidelines and fall prevention policy focus on MVPA, evidence suggests that older adults do not undertake these types of activities frequently. Therefore, guidelines and policy are less relevant to the everyday lives of the people at which they are aimed (Foster & Armstrong, 2018), and opportunities for fall prevention may potentially be missed (Gillespie et al., 2012; Lamb et al., 2011; RCP, 2015; Sherrington et al., 2017). Consequently, a better understanding of how PA, that is not just exercise, affects balance is needed.

1.2.3 Physical activity and the effects on the physiology of older adults

Rudyard Kipling (1902) wrote *"the cure for this ill, is not to sit still, or frowst by a book by the fire, but to grab a large hoe and a shovel also and dig to you gently perspire"*, and research supports that MVPA is a major contributor to successful healthy ageing (PAGAC, 2018). MVPA is associated with the prevention and management of chronic disease (Lee et al., 2012; Newman et al., 2003; Paffenberger et al., 1986; Reiner et al., 2013; Sabia et al., 2012; Umpierre et al., 2011; Wen et al., 2011; WHO, 2010); cognitive capabilities (Angevaren et al., 2010; Baker et al., 2010; Blondell et al., 2014; Fransen et al., 2015; Sofi et al., 2011; Tierney et al., 2010; Zhu et al., 2016), and functional capabilities (Bauman et al., 2016; Chodzko-Zajko et al., 2009; Paterson & Warburton, 2010) (Figure 1.4, page ix).

Additionally, a body of evidence derived from clinical trials suggests that MVPA such as muscle strengthening, balance, and co-ordination exercise can reduce the risk of falls in older adults at higher risk of falling (Bandeen-Roche et al., 2015; Brigola et al., 2015; Bucknix et al, 2015; Cadore et al., 2013; Chen et al., 2014; 2015; Gillespie et al., 2012; Howe et al., 2011; Karlsson et al., 2013; Stubbs et al., 2015; Theou et al., 2011; WHO, 2007). Exercise, as part of a tailored multi-factorial intervention consisting of two or more categories of exercise, or as part of a non-tailored multi-component intervention consisting of, for example, individual risk assessments along with one category of exercise (Howe et al., 2011; Stubbs et al., 2015) either in a home or group setting (Gillespie et al., 2012; Howe et al., 2011; Stubbs et al., 2015), can reverse the effects of ageing on balance and as a result reduce falls in older adults living in institutional care, women, or those with chronic illness (Cadore et al., 2013).

The size of effect between studies varies with a reduced fall rate of between 22% (Relative Risk (RR) 0.78, 95% CI 0.71, 0.86, Karlsson et al., 2013) and 61% (RR 0.39; 95% CI 0.23, 0.66, Stubbs et al., 2015), and a reduced risk of falling of between 13% (RR 0.87; 95% CI 0.81, 0.94, Stubbs et al., 2015) and 35% (RR 0.65, 95% CI 0.51, 0.82, Karlsson et al., 2013). However, most studies focus on higher risk of falling older adults and suggest that higher intensity exercise reduced fall rate (RR 0.60, 95% CI 0.47 to 0.76), and fall risk (RR 0.54, 95% CI 0.35 to 0.83) (Gillespie et al., 2012). Other systematic reviews have also found similar results (Sherrington et al., 2017).

These variations in effect size may be due to the methodological quality of studies where systematic reviews of clinical trials reported small sample sizes; lack of protocol details or adherence to protocol information; general lack of details of included studies; recall bias of self-reported PA, or fall information; varying participant drop-out rates or lack of drop-out information; lack of clarity of the interpretation of the effect due to different effect summary measures (i.e.RR, OR), PA measurements (METS, steps per minute, intensity), or outcome measures (i.e. falls, different functional tests, survey results); and varying participant characteristics such as morbidity level, and gender bias (Howe et al., 2011; Karlsson et al., 2013; Stubbs et al., 2015).

Furthermore, trials of up to one-year duration suggest that walking or cycling have no effect on fall prevention (Gillespie et al., 2012; Howe et al., 2011; Sherrington et al., 2017). However, research suggests that the effects of PA, especially LPA may require a longer period of study than that usually afforded by clinical trials (Bauman et al.,2016; Hoffmann et al., 2016; Morris & Hardman, 1997) where the most frequent duration of trials is three months (Gillespie et al., 2012; Howe et al., 2011). Also, as highlighted in section 1.2.2 the results from cross-sectional studies indicate that LPA is beneficial to balance but an understanding of the benefits of LPA using methodologically robust studies is lacking (PAGAC, 2018; Schmidt et al., 2017).

In summary, the general health benefits of PA, and specifically MVPA, are well documented, but methodologically robust studies supporting the benefits of PA that is not just MVPA, for fall prevention in older adults is needed. For example, LPA such as walking, may require a longer period of study than that afforded by clinical trials, and the methodological robustness of observational evidence is poor. Additionally, studies exploring the effects of PA in lower risk of falling populations is lacking. Therefore, population studies of ageing such as the Irish Longitudinal Study of Ageing (TILDA), the English Longitudinal Study of Ageing (ELSA) and the Northern Irish Cohort of Longitudinal Ageing (NICOLA) (described in Chapters three, four, and five) may address some of the methodological challenges outlined above. For example, these studies include larger representative samples (8,504; 10,601; and 8,500 participants respectively), of repeated measures (every two years for self-reported measures, and every four years for objective health measures) of the same individuals. As a result, the ecological validity of the findings is increased in comparison to clinical trials of smaller sample sizes and shorter timeframes for outcome assessments (Cooke & Iwashyna, 2013; Smith et al., 2011). Consequently, the use of these datasets can provide an opportunity to generalise and make recommendations that may be more relevant for fall prevention in older adults (Cooke & Iwashyna, 2013).

1.3 The challenges of assessing physical activity and balance in older adults using longitudinal data

1.3.1 Physical activity (PA) assessment

The study of PA levels over time poses several challenges to researchers. Firstly, there is no gold standard method recommended for PA assessment (Bauman et al., 2006; Dishman, Washburn & Schoeller, 2001), and Table 1.1 (chapter 1 page i) identifies some of the measures available.

For example, objective measures such as behavioural observation methods using infrared beams to identify usage of recreational areas along with observed activity by researchers, or energy expenditure methods such as calorimetry are not appropriate for use in large population studies because of the high cost associated with device costs and the time required for implementation. In contrast, self-reported measures (e.g. surveys), and motion sensors (e.g. accelerometers) may be more appropriate for the use in large population studies over time. For example, studies such as the UK Biobank study (n=106,053 participants) using wrist-worn accelerometers show the potential for the use of objective measures in population studies over time due to the decline in device cost and increase in usability (e.g. smaller devices; more user-friendly interfaces) (Doherty et al., 2017). However, whilst the use of objective measures is beginning to emerge, self-reported measures such as questionnaires, interviews and surveys have historically been more widely used due to their ease of administration, low cost, and ability to capture both qualitative and quantitative data (Bauman et al., 2009; Dishman et al., 2001; Welk, 2002). (Table 1.2, chapter 1, page ii).

Despite advantages such as low cost, usability, and user acceptability, both motion sensors and questionnaire methods are subject to bias and caution is needed in interpretation of the results. For example, motion sensors such as accelerometers, whilst providing more accurate data relating to actual activity level, can alter the usual activity of participants by motivating higher levels of PA through social desirability. They also lack specificity of PA such as PA type (Bauman et al., 2006; Doherty et al., 2017; Dishman et al., 2001; Haskell & Kiernan, 2000). Also, self-reported measures such as questionnaires, whilst widely used in longitudinal studies due to their ability to measure behavioural change in an older adult population (Shephard, 2003), may be influenced by health status, mood, depression, anxiety, or cognitive ability, as well as seasonal variation in PA patterns, social desirability, or recall issues (Dyrstad et al., 2014; Murphy, 2009; Saelens et al., 2012). As a result, actual levels of activity may be inaccurate, where VPA may be overestimated, and LPA or MPA underestimated (Dyrstad et al., 2014; Saelens et al., 2012). This is of importance in an older population, as older adults tend to spend more time carrying out LPA to MPA (Bauman et al., 2009; 2010; Murtagh et al., 2015), and so methods to minimise these limitations should be considered (Bauman et al., 2009).

1.3.2 Balance assessment in older adults

Research supports the proposal that balance is a complex activity achieved through the integration of multiple inputs from neuromuscular, sensory, and cognitive body systems in the central nervous system (CNS), that elicits an appropriate motor response from the body to maintain or restore postural alignment whilst sitting or standing, moving between postures, or recovering from a slip or trip (Horak, 1995; 2006; Horak & Macpherson, 1996; Lin & Whitney, 2012; Mancini & Horak, 2010; Thomas et al., 2014; Salter et al., 2005) (Figure 1.1, page i).

Balance assessment requires that each system underpinning balance is assessed to enable more appropriate fall prevention programs and effective interventions to be developed (Horak, 1995; 2009; Mancini & Horak, 2010; Sibley et al., 2015). However, there is no standardised method available to measure balance activity, and each type of balance measure presents a challenge for balance screening (Tyson & Connell, 2009).

Firstly, direct measures of balance such as posturography (Jacobs et al., 2006) quantify the position of the body’s centre of mass (COM) in relation to the base of support (BoS) using force platform indicators such as centre of pressure (COP), sway, anterior/posterior or medio-lateral stability, or limits of stability (Winter, 1995). Research suggests that these methods provide a comprehensive assessment for balance, as they remove issues such as variability in test performance; remove subjectivity as they are measured objectively rather than by using a scoring system; and are sensitive to small changes (Visser et al., 2008).

Whilst these measures provide a reliable measure (Mancini & Horak, 2010), they also require expensive equipment more commonly used in a controlled laboratory (e.g. force platforms), and a level of training to use the equipment and interpret the results (Graham, et al., 2008; Tyson & Connell, 2009; Winter, 1995). Thus, their use may be limited to research or within a laboratory setting, rather than clinical practice. Furthermore, whilst direct measures such as wearable inertial sensors with wireless data transfer, such as accelerometers, may address issues relating to cost and improve the mobility of the solutions (Bonato, 2005), challenges with human-device interfacing; quality of the range of measurements; the generation of clinical data based on physiological data; and the visualisation of guides to identify increased risks are still in development for balance assessment (Mancini & Horak, 2010; Rodgers et al., 2014).

In contrast, indirect measures of balance which include for example, observation, self-reporting, objective functional tests, or ordinal scales are less expensive, and may be more appropriate within a clinical or community setting (Graham et al., 2008; Howe et al., 2011; Tyson & Connell, 2009). For example, proxy or functional tests are quick to complete with an average two to three minutes, do not require specialist equipment, can be incorporated into a clinical assessment, and provide meaningful results for both clinicians and participants (Tyson and Connell, 2009; Vieira et al., 2016). Consequently, indirect measures are widely used within clinical practice, where Howe et al.'s (2011) systematic review (n=95 studies) identified approximately 25 different balance measures, and Sibley et al.'s (2015) more recent systematic review (n=66 studies) identified 66 indirect measures.

This multitude of different indirect measures makes it difficult to compare or generalise results across studies. Furthermore, proxy or indirect measures do not individually provide a comprehensive assessment of the factors influencing balance (Horak et al., 2006; Mancini & Horak, 2010; Sibley et al., 2015). Multiple indirect measures can provide a comprehensive assessment of balance if they adopt a systems approach where indirect measures assess neuromuscular, cognitive, and sensory systems collectively (Horak et al., 2006; Sibley et al., 2015). However, there is currently no guidance regarding which combination of individual indirect measures are more effective. Therefore, if a systems approach using indirect measures of balance is to be beneficial for fall prevention, then further research is needed to identify what combination of measures are appropriate (Vieira et al., 2016).

1.3.3 Addressing the issues of assessment using a Structural Equation Modelling (SEM) approach to analysis

Structural Equation Modelling (SEM) is a multivariate approach that integrates different techniques such as between-group and within-group variance comparisons (Analysis of covariance, ANCOVA); path analysis (regression analysis) where equations representing the effect of one or more variables on others can be solved to estimate their relationships; factor analysis or latent class analysis where composite measures such as depression, intelligence, or balance are calculated from observed or measured variables; and latent growth modelling (LGM) to explore the pattern of change over time. SEM has been widely used in social sciences and psychology disciplines but to a lesser extent in medical research and epidemiology (Arlinghaus et al., 2012; Beran and Violato, 2010; Stephenson, Holbert and Zimmerman, 2006; Tu, 2009). It presents several advantages for exploring the association between PA and balance in older adults within this study (Table 1.3 highlights the advantages and limitations of a SEM approach for this study which are further discussed in Chapter three; page iii).

Firstly, measures of PA are subject to bias as outlined in section 1.3.1, and a latent class analysis (LCA), within a SEM approach, enables the calculation of a composite measure of PA intensity using information such as frequency and type of PA and also accounts for measurement error (Bauman et al., 2006; Hagenaars & McCutcheon, 2002; McCoach, Clark & O'Connell, 2007). For example, LCA uses a probabilistic model to classify individuals into mutually exclusive and exhaustive classes based on the pattern of responses in the actual data and removes the shared variance among observed variables to provide a more reliable composite measure of PA (Bauman et al., 2006; Hagenaars & McCutcheon, 2002; Kongstead & Nielsen, 2017; Lanza & Rhoades, 2013; Muthén & Muthén, 1998-2017; McCoach et al., 2007). Therefore, differences between sub-groups or classes become easier to detect, allowing a better prediction of treatment responses or more tailored and effective interventions to be developed (Lanza & Rhoades, 2013; Maslovskaya, Smith & Padmadas, 2018). Additionally, a SEM approach can model both unobserved and observed variables within a measurement model, thus enabling the composite measure of balance to be developed that accounts for the contribution from multiple systems

(Figure 1.1, page vi) (McCoach et al., 2007). For example, as outlined above (section 1.3.2), indirect functional measures of balance are appropriate for use in a clinical and community setting, but guidance relating to which combination of measures across neuromuscular, cognitive, and sensory body systems are appropriate is lacking. Within a SEM approach, multiple observed variables or indirect functional measures (e.g. single legged stance test; gait speed test; handgrip test) can be used to develop the composite measure of balance. Furthermore, confirmatory factor analysis (CFA) (configural, metric, and scalar invariance) can be used to confirm whether the multiple observed measures (indirect measures of balance) can be attributed to the single measure of balance; to assess the reliability of each measure for use within the causal model; and the variance of each observed measure. As a result, unreliability within the composite measure of balance is corrected, so addressing measurement error and increasing the robustness of the findings (Bollen, 1989; McCoach et al., 2007; Raykov & Marcoulides, 2000).

In summary, there is no gold standard for PA assessment and results are subject to bias whether objective or subjective measures are used. A SEM approach can remove measurement error and where self-reported measures are used, can more accurately summarise information from categorical data using an LCA approach. Furthermore, indirect balance measures are more appropriate for a clinical and community setting, but guidance on the combination of measures to comprehensively measure balance is lacking. A SEM approach can address the challenge of combining indirect measures of balance from the multiple systems that contribute to balance, because it deals with latent or unobserved constructs and can also minimise measurement error.

1.4 Summary

Fall prevention is an important health concern for healthy ageing. Global and national PA guidelines and fall prevention policy recognise the importance of MVPA. However, guidelines are lacking for lower risk of falling older adults. Furthermore, adherence to PA guidelines is low, and older adults are more likely to be engaged in LPA rather than MVPA, so guidelines are less relevant. Also, whilst evidence supports the benefits of LPA for wider health issues, robust evidence supporting LPA for balance performance is lacking, and clinical evidence supporting PA for balance performance does not provide an understanding of the longer-term effects (>24 months) of exercise such as walking that may require a longer period of study. Population studies (e.g. TILDA, ELSA and NICOLA) provide an opportunity to explore PA and balance to improve the robustness of the findings. Additionally, a SEM approach provides a method of addressing the challenges of, for example, combining indirect measures of balance within a predictive model of balance, and of addressing the bias from using subjective PA measures.

1.4 Aims and objectives of the thesis

The overall aim of this thesis was to contribute to the understanding of the nature of the relationship between PA and balance in older adults (≥50 years) to inform approaches for fall prevention. Older adults in the context of this thesis are ≥50 years, because this is the age at which body systems required for balance, such as the neuromuscular, cognitive, and sensory systems, have started to decline (Chodzko-Zajko et al., 2009).

This thesis is divided into two strands. Firstly, Smith et al. (2011) suggest that where secondary data analysis is being conducted then a systematic review of the literature should be carried out to help define and frame the research topic. Therefore, the first strand (Chapter two) investigated the existing evidence base to understand the characteristics of PA and how it affects balance in older adults at lower risk of falling. The second strand (Chapters three-six) is informed by the findings from the systematic review (Chapter two) and investigated the effects of PA in older adults using longitudinal data from studies of TILDA, ELSA, and NICOLA to identify an appropriate and robust model of PA and balance. Furthermore, it is recommended that any assumptions identified through secondary data analysis is tested to evaluate the validity of the findings (Craig et al., 2012), and so the model developed was tested using longitudinal data from ELSA and NICOLA. This testing also ensured that the predictive models were relevant to UK and Irish populations. The following objectives are highlighted to achieve the overall aim and in line with the strands highlighted above:

1. To review the evidence base for PA and balance to determine whether PA, that is not just exercise, is associated with improved or maintained balance in older adults at low risk of falling (≥50 years); to identify gaps in the existing evidence; and highlight future research needs in a systematic way.
2. To carry out robust secondary data analysis of longitudinal data to develop a predictive model of PA and balance in older adults (≥50 years) to understand if and how the relationship between PA and balance changes over time using TILDA and ELSA data (Figure 1.5, page x).
3. To test the robustness of the predictive model of PA and balance.

1.6 Thesis overview

A diagram of how the studies are linked to the objectives is presented in Figure 1.5 (page x)

Chapter two describes a systematic review that investigates the association between PA and balance in older adults at low risk of falling (McMullan et al., 2018; published systematic review).

Chapter three explores the relationship between PA and balance using longitudinal data from TILDA.

Chapter four develops the composite measure of PA for use in Chapter five. The composite measure of PA combines information relating to PA type and frequency provided by the categorical self-reported PA measure used in the ELSA study.

Chapter five explores the trajectory of change between PA and balance using the measure of PA developed in Chapter four.

Chapter six assesses the robustness of the model of balance developed using TILDA, ELSA, and NICOLA data.

Chapter seven summarises the findings and discusses some of the points raised in this thesis, along with the implications for clinical practice and future research.

Table 1.1
A summary of physical activity assessment methods

Method	Examples	Population study? (yes/no)	Setting	Cost (High/low)	Activity type measure? (Yes/no)	Frequency measure? (Yes/no)	Duration measure? (Yes/no)	Intensity measure? (Yes/no)	Other issues
Behaviour observation	Motion sensors (e.g. infrared beam)	No	Defined space e.g. park	High cost of devices & time for observation	Yes	Yes	Yes	No, but could be estimated	Electronic surveillance may not be socially acceptable. Can alter usual activity
Motion sensors	Pedometer; accelerometer	Yes	Clinical, home	Cost has declined & use in large population studies increased	No	Yes	Yes	Yes (accelerom eter)	Can alter usual behaviour i.e. act as a motivator therefore increasing usual activity
Measures of energy expenditure	Calorimetry; doubly labelled water; heart rate	No	Laboratory	high	No	No	No	Yes	Can interfere with usual activity
Self-reported methods	IPAQ	Yes	Clinical or laboratory	low	Yes	Yes	Yes	Not directly but can be estimated	Recall & Interpretation bias

(supporting references: Bauman et al., 2006; Doherty et al., 2017; Dishman et al., 2001; Haskell & Kiernan, 2000)

Table 1.2

Advantages and disadvantages of using self-reported measures for the assessment of physical activity in older adults (≥50 years)

Self-reported instrument	Reference	Description	Advantages for use within an older adult population	Limitations for use within an older population
International Physical Activity Questionnaire (IPAQ)	Craig et al., 2003.	-Recall: last 7 days. -Short form: frequency, duration of time spent on walking/ vigorous/ moderate/ sedentary activity.	-Lower intensity activity is assessed. -Last 7-day recall limits recall bias. -Short form available. -Recommended use for standardisation for comparisons between populations.	Last 7-day recall bias.
Communities Health Activities Models Program for Seniors (CHAMPS)	Harada et al., 2001.	-Recall: a typical week in past month. -Weekly frequency, duration and intensity of PA. -leisure time, housework, and social activities.	-Assesses usual activities. -Developed for older adults living in the community.	-Last 7-day recall bias. -Definition of a typical week may be different for different people.
Yale Physical Activity Survey (YPAS)	DiPietro et al., 1993.	-Recall: a typical week in last month and activities during the past month (e.g. walking and standing). Participant is shown 28 different types of PA.	-Assesses usual activities. -Interview administered questionnaire. -Uses visualisation so increases recall.	-Last 7-day recall bias. -Definition of a typical week may be different for different people.

(adapted from Bauman et al., 2006)

Table 1.3
Advantages and limitations of the role of using a structural equation modelling (SEM) approach for the development of a predictive model of physical activity and balance in older adults

Requirement	SEM advantage	Limitations	Other analytical approaches	Conclusion
To develop a composite measure of balance. To develop a composite measure of PA.	SEM distinguishes between latent variables which are constructs that are not directly observed (e.g. balance) and observed variables (e.g. those variables that are directly observed such as the individual indirect or proxy measures for measures of balance). Confirmatory Factor Analysis can combine multiple observed measures and address measurement error. A latent class approach can develop a composite measure that is based on probabilities rather than arbitrary cut-off points.	- Single measures or indicators may only be available therefore not representative of the latent variable. This can be resolved using ‘parcels’ (subsets of the summative variable e.g. CSED summative score parcelled into depression, anxiety, and hopefulness). -Omitted variables may bias the results of the measurement or structural model.	Latent measures are not applied to any other analysis method.	-SEM includes both observed and latent variables in SEM allows a wider variety of hypotheses to be tested. For example, balance is not directly measured, but observed variables across neuromuscular, cognitive and sensory systems provide a proxy measure. -PA may be self-reported categorical measures and SEM provides the ability to construct a latent PA variable. For example, data such as PA type and frequency information, LCA can develop a latent construct of PA intensity.

(Supported by Beran & Violato, 2010; Hox & Stoel, 2005; Schafer & Graham, 2002; Stephenson et al., 2006; Tomarken & Waller, 2005

Table 1.3 (continued)

Advantages and limitations of the role of a Structural Equation Modelling (SEM) approach for the development of a predictive model of physical activity and balance in older adults

Requirement	SEM advantage	SEM Limitations	Other analytical approaches	Conclusion
There is no guidance on which measures comprehensively measure balance. Measurement error of self-reported measures such as PA need to be addressed to improve the robustness of findings.	SEM removes measurement error. For example, SEM uses multiple indicators to estimate the effects of the latent variable and through confirmatory factor analysis simultaneously with path analysis addresses unreliability within the construct.	n/a	Measurement error is not addressed in other analytical approaches and can reduce the regression weight from the predictor to the dependent variable leading to underestimation of the effect and biased results.	SEM removes measurement error therefore addressing the challenges posed by using indirect measures of balance and self-reported measures of PA.
The development of a model of PA and balance requires the model of balance using indirect or proxy measures to be developed, bias caused by measurement error to be removed, as well as testing the relationship between both PA and balance.	-SEM is a framework of multiple models e.g. ANCOVA, multiple regression, path analysis, confirmatory factor analysis, latent growth. -SEM can explore the relationship between dependent variables, and between dependent and independent variables. -SEM can concurrently test models (e.g. the measurement model (the degree of correlation between the observed variables) and structural model (the degree of correlation between the latent variables). -SEM can calculate direct, indirect and total effect because more than one exogenous and endogenous variables are estimated simultaneously.	-Errors can result if variables are omitted. -Multiple models complicate results and caution should be taken in the interpretation of the results.	-ANCOVA, MANCOVA can only show a single relationship between independent and dependent variables. -Equations are solved one by one using separate models – compartmentalised	SEM provides a framework that enables a holistic approach to analysis and allows a wider variety of hypotheses to be tested.

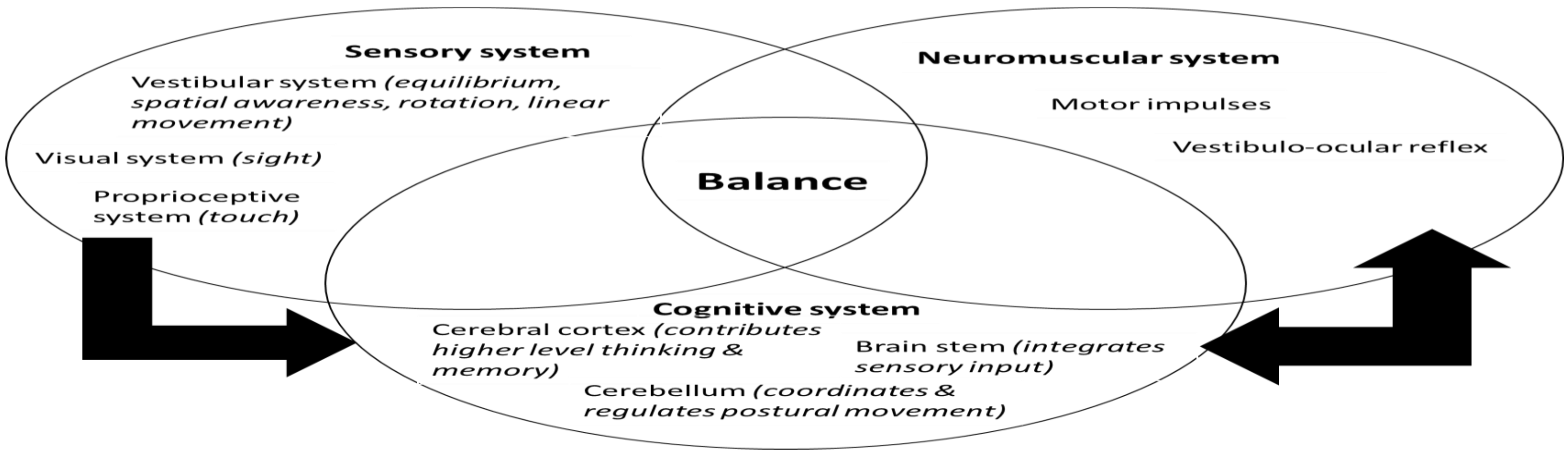
(Supported by Beran & Violato, 2010; Hox & Stoel, 2005; Schafer & Graham, 2002; Stephenson et al., 2006; Tomarken & Waller, 2005)

.

Table 1.3 (continued)
Advantages and limitations of the role of a Structural Equation Modelling (SEM) approach for the development of a predictive model of physical activity and balance in older adults

Requirement	SEM advantage	SEM Limitations	Other analytical approaches	Conclusion
How do we assess the model of PA and balance to ensure it is appropriate?	-SEM uses a variety of fit statistics (outlined in Chapter 3) to systematically assess model fit. -SEM enables group comparisons allowing models to be validated using different datasets.	-There is no gold standard regarding which fit statistics should be used so the process is open to manipulation. -Omitted variables can produce biased models.	No systematic evaluation of the theoretical model.	-Whilst no guidance on which fit statistics are most appropriate, they still identify the best model fit systematically. -SEM can be used to compare the model of PA and balance using different datasets for validation.
There is a need for longitudinal analysis to understand the association between PA and balance, but these studies may be prone to missing data due to the time period over which it is gathered (drop outs, morbidity)	Although not unique to SEM, a robust form of Maximum Likelihood Estimation (MLE) is a model-based estimation strategy for missing data reducing standard errors.	SEM results are subject to sampling effects in respect to individuals, measures, and occasions, but cross validation can improve generalisability.	Methods such as listwise deletion, pairwise deletion and mean imputation use summative statistics that can produce biased and inefficient parameter estimates, inaccurate standard errors, confidence intervals with poor coverage probabilities, invalid hypothesis tests	Missing data in longitudinal data is an issue and SEM's robust MLE addresses the issues with more traditional methods.
Sample size	Sample size 20-25 times number of parameters to be estimated or at least 2000.	Vague sample size guidance.	Smaller sample sizes.	Sample size specific.

(Supported by Beran & Violato, 2010; Hox & Stoel, 2005; Schafer & Graham, 2002; Stephenson et al., 2006; Tomarken & Waller, 2005)



Sensory information is sent to the brain such as motion, equilibrium, and spatial orientation information provided by the vestibular apparatus in each ear; visual cues identifying how a person is orientated to other objectives is provided by sensory receptors in the retina (rods and cones); and proprioceptive information from the sensory receptors in the skin, muscles and joints that are sensitive to stretch or pressure in the surrounding tissues. This sensory information is then sorted and integrated with learned information contributed by the cerebellum (such as automatic movements that have been learned through repeated exposure to certain motions e.g. hitting a tennis ball) and the cerebral cortex (such as previously learned information such as how to navigate a slippery surface). As sensory integration takes place, the brain stem transmits impulses to the muscles that control movement of the eyes, head, neck, trunk, and legs therefore allowing a person to both maintain balance and have clear vision (Myers et al., 1996, 1998; Shumway-Cook & Woollacott, 2001; Sibley et al 2015; Thomas et al., 2014).

Figure 1.1
The balance system

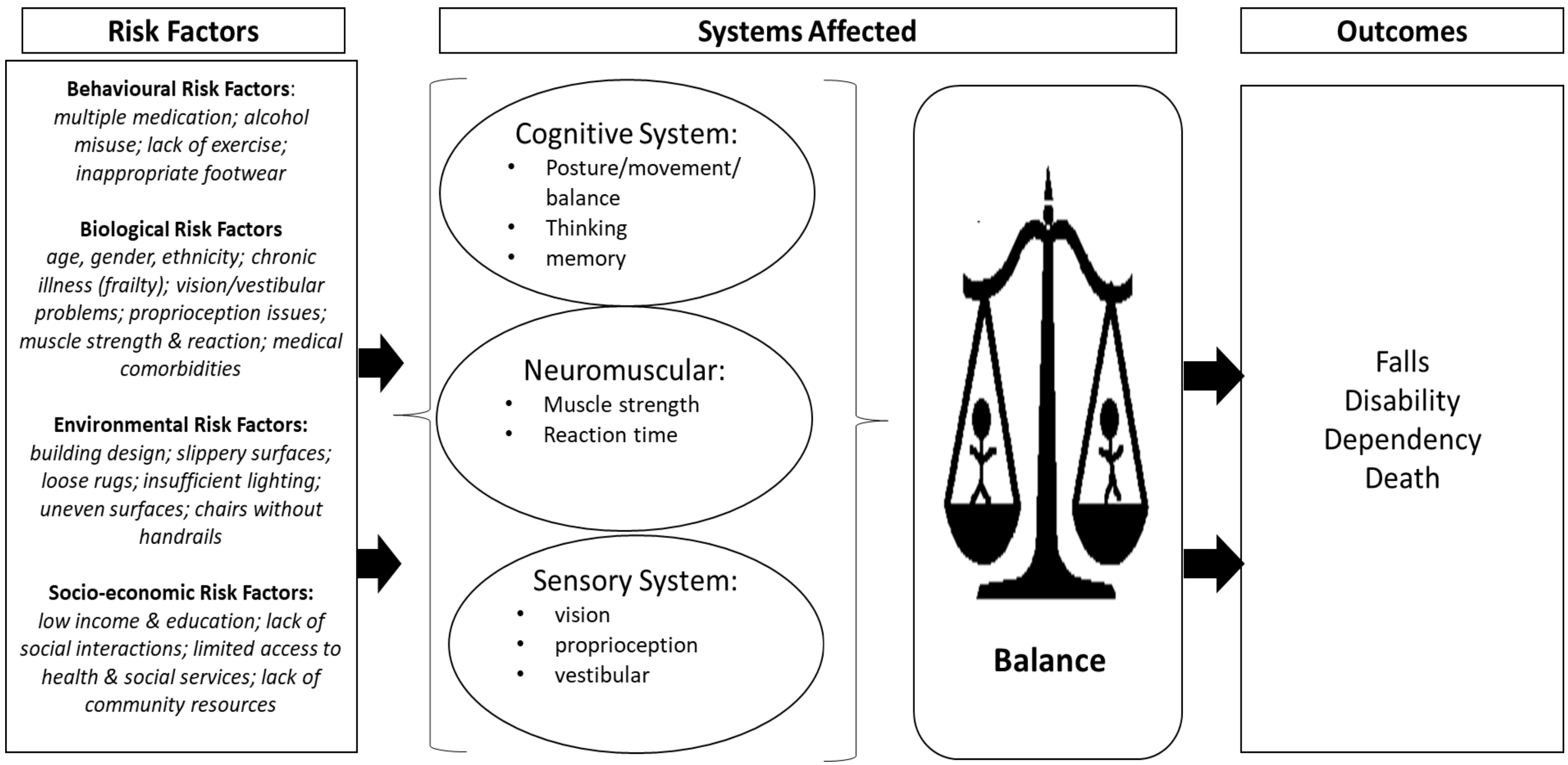
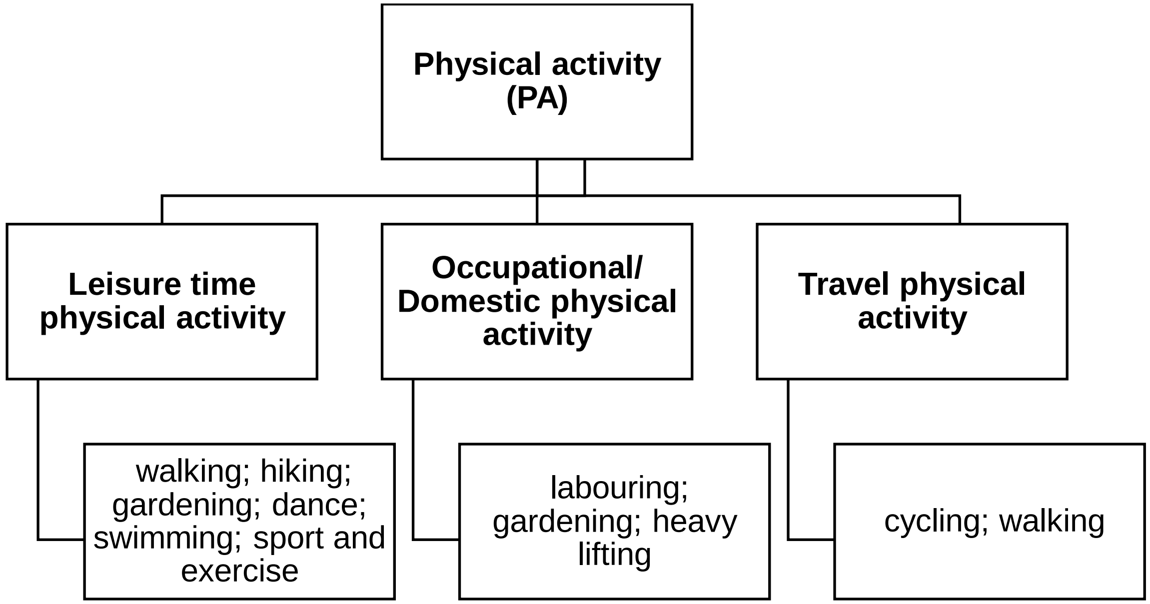
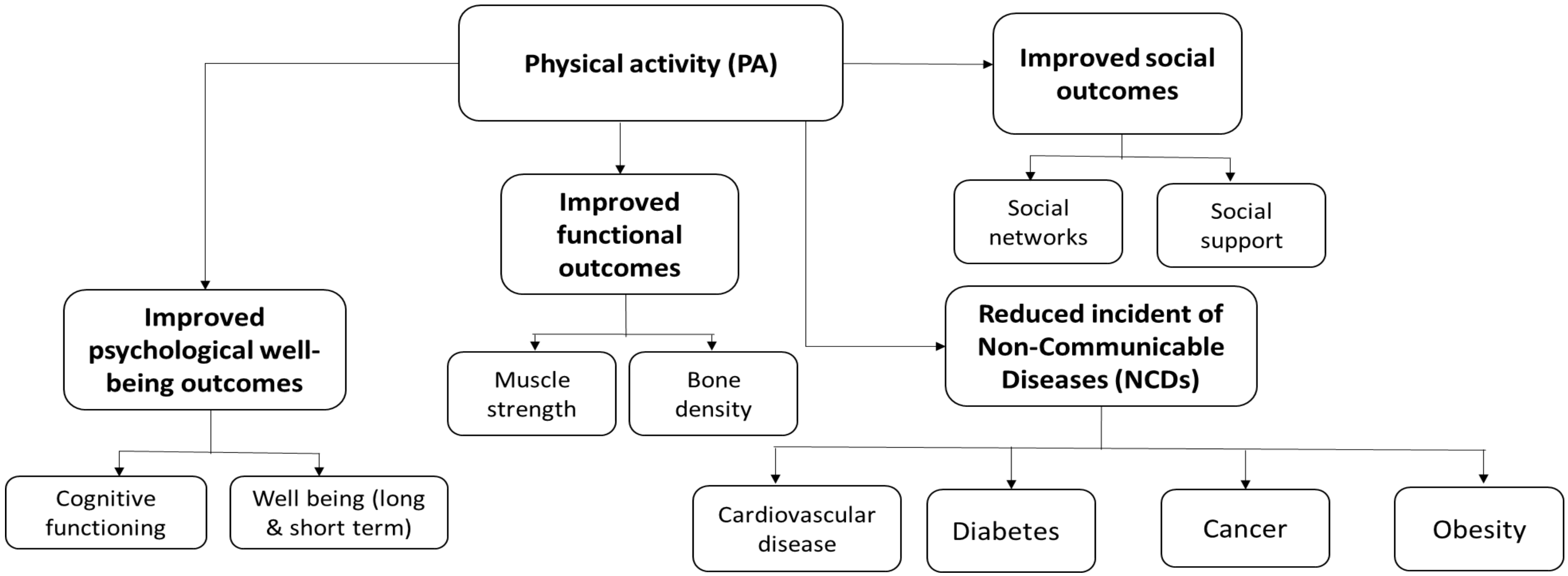


Figure 1.2
A summary of risk factors affecting balance in older adults (≥50 years)



(supporting references: Howley, 2001; WHO, 2011; 2013)

Figure 1.3
Physical activity description characterised by domain



(supporting references: Newman et al., 2003; Bauman et al., 2016; Demakakos et al., 2010; Paterson & Warburton, 2010; Reiner et al., 2013; Sabia et al., 2012; Savela et al., 2013; Valliant & Mukamai, 2006).

Figure 1.4
The benefits of physical activity to health in older adults (≥50 years)

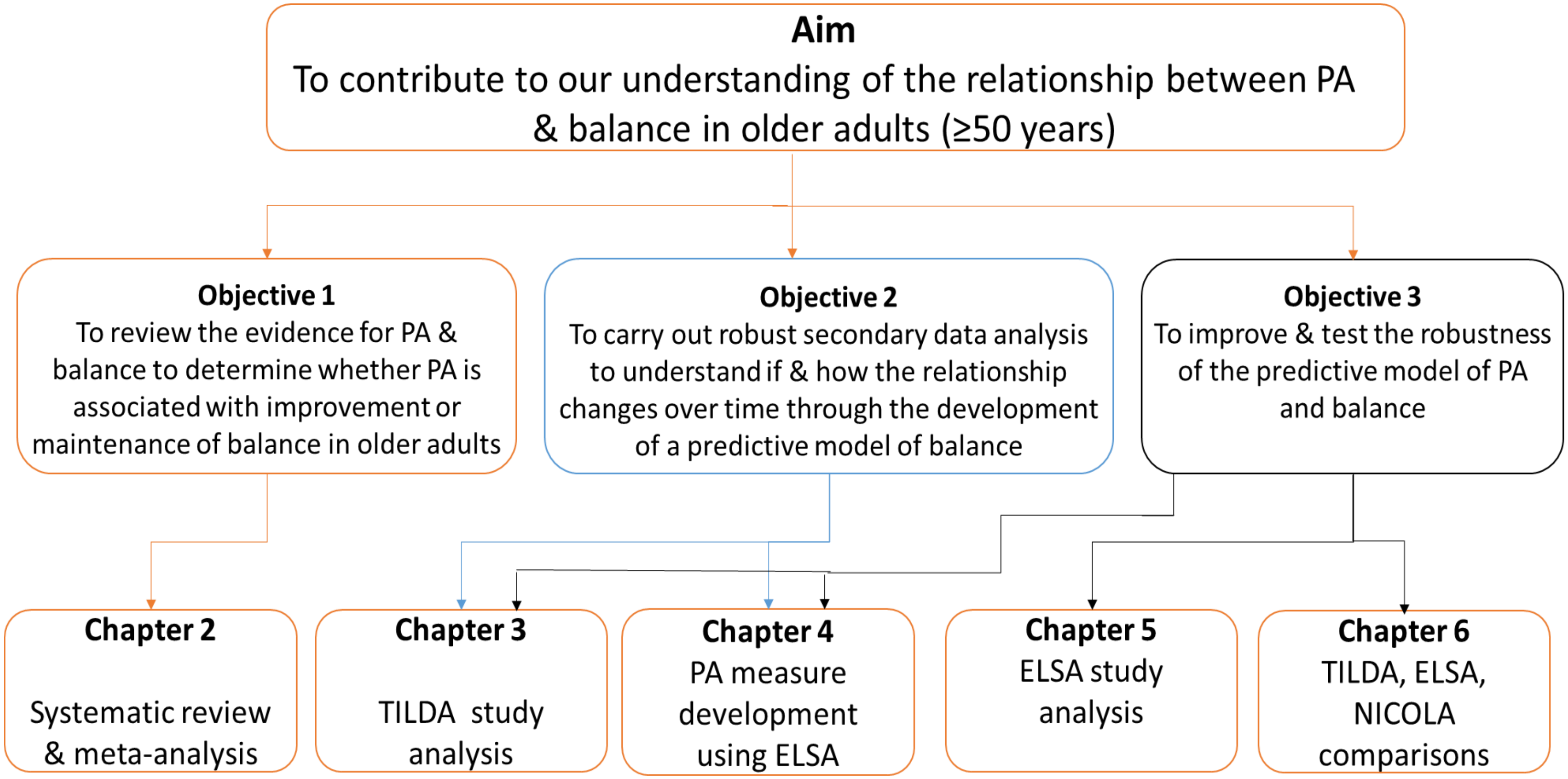


Figure 1.5
Schematic diagram of how the chapters meet the overall aims and objectives of the thesis

Chapter Two:

The association between balance and free-living physical activity (PA) in an older healthy community-dwelling adult population (≥50 years): A systematic review and meta-analysis

(This Chapter is published: McMullan II, McDonough SM, Tully MA, Cupples M, Casson K & Bunting BP. (2018). The association between balance and free-living physical activity in an older healthy community-dwelling adult population: A systematic review and meta-analysis. BMC Public Health, 18, 431. Doi: https://doi.org/10.1186/s12889-018-5265-4)

2.1 Abstract

Aim

Poor balance is associated with an increased risk of falling, disability and death in older adults, and fall prevention is an important health consideration. The benefits of exercise, a structured and planned subgroup of physical activity, is well understood. However, less is understood about the effects of PA, that is not exercise, on balance, and guidelines regarding healthy older adults are vague. Therefore, this review explores the effects of PA, that is not just exercise, on balance in healthy older (≥50 years) community-dwelling adults to inform fall prevention policy, and as a result reduce the human and economic cost of falls.

Methods

Electronic literature searches of relevant databases (CENTRAL, Bone, Joint and Muscle Trauma Group Specialised register and CDSR in the Cochrane Library, MEDLINE, EMBASE, CINAHL, PsychINFO, and AMED) were searched from inception to 7th June 2016. The inclusion criteria were intervention and observational studies investigating the effects of free-living PA on balance in healthy community-dwelling adults (≥50 years). Thirty studies were eligible for inclusion. Data extraction and risk of bias assessment were independently carried out by two review authors. Due to the level of heterogeneity in studies, balance outcomes from observational studies were pooled as standardised mean differences (SMD) and 95% confidence intervals (CI), and outcomes from RCTs were synthesised using a best evidence approach.

Results

Limited evidence provided by a small number of RCTs, and evidence from observational studies of moderate methodological quality suggest that free-living PA habitually carried out of between one- and 21-years' duration improves measures of balance in older healthy community-dwelling adults. Statistical analysis of observational studies found significant effects in favour of more active groups for neuromuscular measures such as gait speed (SMD 0.66 m/s; 95% CI 0.26 to 1.06 m/s, 194 participants, two studies); functionality using timed up and go (TUG) (SMD -0.70 s; 95% CI -1.03 to -0.37 s; 161 participants, two studies), Single Leg Stance (SLS) (SMD 1.17s; 95% CI 0.74 to 1.60s, 110 participants, two studies), and Advanced Balance Confidence (ABC) (SMD 1.47; 95% CI 0.70 to 2.25, 155 participants, three studies); flexibility using forward reach test (SMD 0.80m; 95% CI 0.48 to 1.11 m, 193 participants, two studies); strength using isometric knee extension (SMD 0.64, 95% CI 0.35 to 0.94, 292 participants, three studies) and ultrasound (SMD 0.57; 95% CI 0.25 to 0.89, 158 participants, two studies). A significant effect was also observed for less active groups on a single sensory measure of balance, the knee joint repositioning test (SMD -1.37; 95% CI -2.29 to -0.45, 58 participants, two studies).

Conclusion

There is some evidence that free-living PA is effective in improving balance outcomes in older healthy community-dwelling adults, but future research should include higher quality studies that focus on a consensus of balance measures that are clinically relevant and explore the effects of free-living PA on balance over the longer-term.

2.2 Introduction

Ageing is a natural human condition, but falling should not be considered an inevitable part of the ageing process, and a body of evidence derived from clinical trials suggests that exercise, a sub-category of PA that is structured, planned, repetitive, and carried out over a relatively short time frame (one to 12 months with the most frequent being three months), can maintain or improve balance in older adults at high risk of falling (e.g. those with chronic disease; living in institutional care; or women) (Bucknix et al., 2015; Cadore et al., 2013; Gillespie et al., 2012; Howe et al., 2011; Karlsson et al., 2013; Stubbs et al., 2015). Additionally, as highlighted in Chapter one (section 1.2.2) PA guidelines support MVPA for general health benefits within the general population, as well as balance and co-ordination activities for older adults at higher risk of falling (CMO, 2011; DoH), 2011; Sherrington et al., 2017; WHO, 2010).

However, whilst research and guidelines support the benefits of MVPA, older adults are more likely to engage in LPA (Arnardottir et al., 2013; Bann et al., 2015) due to health-related issues or physical capability (Ainsworth et al., 2011; Schutzer & Grave, 2004). Additionally, evidence supporting LPA such as walking for leisure, light housework, or light gardening activities for balance is lacking perhaps because guidelines are based on the available evidence that suggests that MVPA is beneficial for health in older adults (Bauman et al., 2016; Milton et al., 2018; NICE, 2013). Furthermore, guidance is less explicit in terms of PA for balance performance for older adults at lower risk of falling (NICE, 2013; PHE, 2017; WHO, 2007; 2010).

Therefore, opportunities for fall prevention are potentially being missed (Gillespie et al., 2012; Lamb et al., 2005; RCP, 2015; Sherrington et al., 2017), and a better understanding of the effects of free-living PA for the purposes of leisure, occupation, and travel is needed in healthy older adults (Cooper et al., 2010; 2011; PAGAC, 2018; Reiner et al.,2013).

2.3 Aims and objectives

The primary aim of this systematic review was to investigate the effect of free-living PA, defined as leisure activity based on personal interests and needs (walking, hiking, gardening, swimming, sport, and dance), travel activity (cycling or walking), occupational activity (labouring, gardening, heavy lifting), or planned exercise in the context of daily, family, and community activities (walking programmes, swimming clubs, Tai Chi clubs) (Howley, 2001; WHO, 2011 & 2013), on balance in older adults (≥50 years) who are at low risk of falling. Two specific objectives of this review were:

1. To assess whether free-living PA is effective in improving or maintaining balance in older adults (≥50 years) who are at lower risk of falling.
2. To identify key characteristics of free-living PA that are associated with greater benefits to balance in older adults at lower risk of falling (≥50 years).

2.4 Methods

2.4.1 Study design

This review followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework (Liberati et al., 2009) and the Cochrane Handbook for Systematic Reviews of Interventions (Higgins & Green, 2011). A protocol was developed *a priori* using the same guidelines to strengthen the methodological approach and was registered on PROSPERO (CRD42016039114). The changes made between the protocol and the systematic review are documented in Appendix I.

2.4.2 Criteria for considering studies

2.4.2.1 Types of studies

Studies were included if they used an intervention design (regardless of randomisation procedures), or an observational design; included a comparison group; were published in English; were peer-reviewed; and had full text available. Case series, case reports, cross-over trials and retrospective case-control studies were excluded due to the potential higher risk of bias in these designs (Higgins & Green, 2011).

2.4.2.2 Types of participants

Studies involving healthy community-dwelling adults ≥50 years were included.

Unhealthy older adults suffering or recovering from conditions that might impact on balance were excluded. For example, those suffering from neurological conditions (e.g. Parkinson's disease, multiple sclerosis, dementia, Alzheimer's) (Gillespie et al., 2012); somatosensory disorders (e.g. vision or hearing impairment (Hoffmann et al., 2016); amputation of lower or upper limbs, musculoskeletal disease (e.g. osteoporosis, arthritis, or muscular disease), or those with a history of falls or fractures (Howe et al., 2011) were excluded.

2.4.2.3 Types of intervention

PA aimed at improving or maintaining balance in older adults is considered in this review, regardless of setting (e.g. home or community based), implementation (e.g. by a health professional or individual), or delivery method (e.g. face to face or telephone). A high number of clinical trials have focused on structured and planned exercise interventions of a relatively short duration (Gillespie et al., 2012; Howe et al., 2011; Stubbs et al., 2015), and as a result, studies were included that met the definition of free-living PA that includes leisure, travel, occupational, and planned exercise in the context of daily, family, and community (Howley, 2001; WHO, 2011; 2013), but excluded structured and/or planned exercise that took place in a researcher environment or a healthcare facility.

In addition, research suggests that PA should reduce the need for upper limb support (Sherrington et al., 2017), and seated PA (e.g. Tai Chi while sitting) may not improve standing or dynamic movement related balance as the neuromuscular system is not challenged adequately to improve balance (Rose, 2005). Thus, studies that included only seated or upper limb PA were excluded. Additionally, studies that included only interventions such as drug therapy, supplements such as Vitamin D, or educational or counselling programmes were excluded.

2.4.2.4 Types of outcome measures

The main outcome of interest was balance, so studies reporting a measure of balance involving neuromuscular, cognitive or sensory systems, either as a primary or secondary outcome were included in accordance with Horak's (2006) theory of balance (Chapter one).

Primary outcomes

Primary outcomes of interest focused on indirect measures of balance as research suggests that tests should be easy to administer, not require laboratory equipment, and easy to interpret to be successfully implemented (Graham et al., 2008). Indirect measures of balance such as, but not limited to, neuromuscular measures of strength, gait, functionality, and flexibility (e.g. grip test (Bohannon & Schaubert, 2005); chair stand test (Whitney et al., 2005); walking speed test (Lienhard et al., 2013); Timed Up & Go test (TUG) (Podsialdo & Richardson, 1991); functional reach test (Toshiaki et al., 2006); Activities of Balance Confidence (ABC) (Powell & Myers, 1995); cognitive measures (e.g. Mini Mental State Exam (MMSE) ; Folstein, Folstein &

McHugh, 1975); sensory measures (e.g. knee joint repositioning); and adverse events such as fall rate or number (Gillespie et al., 2012).

Secondary outcomes

Direct measures of balance included force platform indicators such as centre of pressure (COP), sway, anterior/posterior or medio-lateral stability, or limits of stability (Winter, 1995).

2.4.3 Search strategy

2.4.3.1 Electronic databases

The following eight databases were searched: the Central Register of Controlled Trials (CENTRAL), the Cochrane Database of Systematic Reviews (CDSR), and the Cochrane Bone, Joint and Muscle Trauma Group Specialised Register in the Cochrane Library (June 2016); Ovid Medical Literature Analysis and Retrieval System Online (MEDLINE) (R) (1946 to 31st May 2016), Ovid MEDLINE®—includes new records, not yet fully indexed; Ovid EMBASE (1974 to 31st May 2016); EBSCO Cumulative Index to Nursing and Allied Health Literature (CINAHL) plus (1937 to 31st May 2016); Ovid PsycINFO (1806 to 31st May 2016); and Ovid Allied and Complementary Medicine Database (AMED) (from 1985 to 31st May 2016). The syntax and Medical Subject Headings (MeSH) of the Medline search strategy were modified for the additional seven databases and is shown in Appendix II.

2.4.3.2 Searching other resources

Systematic reviews that were eligible were screened for additional relevant references. The electronic search strategy was supplemented by using the Science Citation Index to perform citation tracking of the trials identified by the first step. Additionally, to identify other potential studies, the National Institute for Health Research library (www.nihr.ac.uk) was searched, as well as published research on the longitudinal studies of ageing from ELSA (http://www.elsa-project.ac.uk/) and TILDA (http://tilda.tcd.ie/). No further studies were identified based on the inclusion and exclusion criteria for the review.

2.4.4 Data collection and analysis

2.4.4.1 Selection of studies

Results from the searches were imported into REFWorks (2.0), and duplicates removed (IM). The remaining study titles, keywords and abstracts were independently screened by two review authors (IM; SM/MC/KC). Following this, the full text of selected articles was obtained and screened independently by two review authors (IM; SM/MC/KC). All studies excluded at full text screening stage and the reasons for exclusion were recorded. Review authors were not blinded with respect to author name, journal or date of publication during this process. Any discrepancies were resolved by consensus or by third party adjudication (Figure 2.1, page xliii).

2.4.4.2 Assessment of methodological quality

Risk of bias assessment was carried out independently by two reviewers (IM and SM for intervention studies; IM and KC for observational studies) post a trial with a small number of studies to check for understanding. Disagreements were resolved by consensus or third-party adjudication. The Cochrane Collaboration tool (Higgins & Green, 2011) was used to assess the quality of included intervention studies by considering their internal validity and risk of bias (Appendix III). The approach considers studies are low risk of bias where risk is low across all domains or most information was from studies at low risk; unclear risk where risk is unclear across all domains or most information was from studies at unclear risk; and high risk of bias where one or more domains were high risk or the proportion of information from studies at high risk was sufficient to affect the interpretation of the results.

There is no recognised 'gold standard' appraisal tool for risk of bias in observational studies (Sanderson et al., 2007), but guidelines suggest that the Newcastle Ottawa Scale (NOS) provides an appropriate and adequate tool (NICE, 2015; Higgins & Green, 2011; Hartling et al., 2013). Using a variation of the NOS (Herzog et al., 2013; Wells et al., 2010), the risk of bias for observational studies was assessed using the domains of: selection of study groups, sample size, comparability, and ascertainment of exposure and outcome (Appendix IV). In the absence of formal threshold scores for rating quality (Wells et al., 2010) studies were rated as high risk of bias or low quality if scored four stars or below, moderate risk of bias or quality if scored five to seven stars, and low risk of bias or high quality if scored eight stars and above (maximum stars possible was ten).

Where insufficient data were available to complete a meta-analysis, the data were synthesised qualitatively using a best evidence synthesis advocated by van Tulder et al. (2000) where evidence is considered 1) strong; consistent findings in multiple RCTs assessed as having low risk of bias; 2) moderate; consistent findings in one RCT assessed as having low risk of bias, and one or more RCTs assessed as having high risk of bias, or by generally consistent findings in multiple RCTs assessed as having high risk of bias; 3) limited or conflicting evidence; only one RCT (assessed as having either a low or high risk of bias), or inconsistent findings in multiple RCTs; and 4) no available evidence; no published RCTs that have assessed interventional effect. Consistency was defined as 75% or more of studies (van Tulder et al., 2000).

2.4.4.3 Data extraction and management

Two review authors independently extracted data using a pre-tested data extraction form (IM; SM/MT) (Appendix V). Disagreements were resolved by consensus or third-party adjudication.

Where available, the following information was extracted and summarised:

- Methodological characteristics: study design; number of participants; recruitment method; timeframe of follow-up.
- Participant characteristics: gender, age (range, median, mean), education level; health status; setting; BMI; ethnicity.

- PA characteristics: measuring tool; type of PA; PA level or type of comparison group; PA duration, frequency and intensity.
- Outcome characteristics: test used; test validation; results (standardised mean difference (SMD), mean and standard deviation (SD), or Risk Ratio (RR)).

2.4.4.4 Measurement of treatment effect

Continuous outcomes were expressed as the standardised mean difference (SMD), or mean difference (MD) where possible, with 95% confidence intervals (CI). For example, mean and standard deviation (SD) for key balance outcome measures were extracted and the standardised mean difference and 95% confidence intervals or mean difference and 95% confidence intervals were calculated using Cochrane Collaboration Review Manager (Revman) (v 5.1):

$$\text{SMD} = \frac{(\text{mean in more active group} - \text{mean in less active group})}{\text{Pooled SD of both groups}}$$

MD= (mean in more active group – mean in less active group)

SMD or MD was chosen where included studies used varying measures/scales or the same measures/scales to assess balance performance respectively. A random effects model was chosen as data were extracted from a series of studies performed by researchers operating independently and the objective is to generalise the findings (Higgins & Green, 2011).

Dichotomous outcomes were expressed as a risk ratio (Relative Risk - RR) with 95% confidence interval (CI):

$$RR = \frac{\text{Risk of fall in more active group}}{\text{Risk of fall in less active group}}$$

Ordinal outcomes were expressed either using RR or standardised difference in the means (SDM) where appropriate.

2.4.4.5 Unit of analysis issues

Where studies involved multiple intervention groups and more than one group met the inclusion criteria, then to avoid double counting, PA interventions were only compared to minimal intervention controls (O'Connor et al., 2015) in accordance with Ainsworth et al.'s (2000) Compendium of Physical Activities. Where studies included groups that compared PA levels by, for example, gender or age rather than by 'less' or 'more' PA, then where data were available, these groups were combined (Higgins & Green, 2011) (Appendix VI).

2.4.4.6 Missing data

Where missing data were discovered during data extraction the authors were contacted to request the required information. Where data were not provided by authors then, where graphical representation of the data was available, the results were estimated by two reviewers. The potential effect of missing data on the conclusions drawn from this review were considered in the results section.

2.4.4.7 Assessment for heterogeneity

Using RevMan (v. 5.1) statistical heterogeneity was assessed using a combination of visual inspection of the graphs as well as consideration of the chi-squared (χ^2) test, p value ($p<0.1$), and the I^2 statistic where values of up to 40% were considered unimportant, 30% to 60% were moderately important, 50% to 90% were substantially important, and 75% to 100% were considerably important (Higgins & Green, 2011). In cases of significant heterogeneity studies were pooled using a random effects model.

2.4.4.8 Assessing reporting bias

Funnel plots that included effect size and standard error were used to examine asymmetry and to assess reporting bias.

2.4.4.9 Data synthesis

Studies were grouped by study design type (Borenstein et al., 2009) and then according to balance outcome (direct or indirect) (Howe et al., 2011; Winter, 1995). Where data were available and appropriate as per the guidelines outlined by the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins & Green, 2011) a statistical analysis was conducted in RevMan (5.1). A random-effects model was used to pool the analyses due to the statistical and clinical heterogeneity in the balance measures being combined. Where insufficient data were available a best evidence synthesis was completed as advocated by van Tulder et al., (2000).

2.4.4.10 Subgroup analysis and investigation of heterogeneity

Based on the grouping of study design and balance assessment tools, heterogeneity was explored where possible using the following subgroup analyses:

- PA type where the effects of 3D PA that is constant, controlled, fluid, and repetitive movement in forward, back, side to side, and up and down direction (e.g. Tai Chi and Qi Gong) versus general activity PA that is for the purposes of leisure, occupational or travel purposes (e.g. walking, golf, bowling) as defined by the taxonomy of fall prevention classification system (ProFaNE) (Lamb et al., 2005) was explored on balance outcomes.
- Participant age group where the effects of young-old (mean age 50-64yrs) versus old-old (≥65yrs) were explored on balance outcomes.

A random effects model was used to pool data in all subgroup analyses and the I^2 test for heterogeneity across subgroups reported by Revman (5.1) (where values of up to 40% were considered unimportant, 30% to 60% were moderately important, 50% to 90% were substantially important, and 75% to 100% were considerably important; Higgins & Green, 2011), as well as the within-group effect size were used to determine whether there was evidence for a difference in treatment effect between subgroups. Within-group effect size was calculated as:

$$\frac{\text{Mean post treatment – mean pre-treatment}}{\text{SD pre-treatment}}$$

2.4.4.11 Sensitivity analysis

Post-hoc sensitivity analyses were carried out to assess the possible influence of risk of bias and heterogeneity on the robustness and overall validity of the results (Higgins & Green, 2011). For example, observational studies rated as high risk of bias or low quality if scored four stars or below using the NOS and where heterogeneity was high (>60%) were excluded from analysis. If the sensitivity analysis appeared to influence the findings of the review, for example, exclusion of the studies changed the level of significance then this was reported in the 'Results' section.

2.5 Results

2.5.1 Study selection

A total of 2,364 papers were identified by the search strategy. From the title, abstract, and keywords, two reviewers independently identified 82 relevant studies for full text review. From the full text review, 52 were excluded (Table 2.1, page xi) resulting in 30 papers being reviewed (n=1,547 participants). The process, including the reasons for exclusion, is shown in Figure 2.1 (page xIiii) (Liberati et al., 2009).

2.5.2 Observational studies

Table 2.2 (page xiv) shows a summary of the characteristics of included observational studies.

2.5.2.1 Design, sample size, and location

Twenty-six studies were observational (one prospective cohort: Aoyagi et al., 2009, and 25 cross-sectional studies). Sample size ranged from 23 (Gao et al., 2011) to 170 (Aoyagi et al., 2009) with an average of 54 participants, but only one study carried out a sample size calculation (Wayne et al., 2014).

Fourteen studies did not specify study location (Brooke-Wavell & Cooling, 2008; Buatois et al., 2007; Fong & Ng, 2006; Fong et al., 2014; Gauchard et al., 1999; Gauchard et al., 2001; Gauchard et al., 2003; Hakim et al., 2010; Rahal et al., 2015; Tsang & Hui-Chan, 2004; Tsang & Hui-Chan, 2005; Tsang & Hui-Chan, 2006; Tsang & Hui-Chan, 2010; Zhang et al., 2011); one study was carried out in Japan (Aoyagi et al., 2009); four in China (Gao et al., 2011; Gyllensten et al., 2010; Lu et al., 2013; Tsang et al., 2004); two in Taiwan (Wong et al., 2001; Wong et al.,2011); one in the UK (Dewhurst et al., 2014); two in US (Hakim et al., 2004; Wayne et al., 2014); one in Brazil (Gaudagnin et al., 2015); and one in France (Perrin et al., 1999).

2.5.2.2 Participants

Participants across all studies were defined as healthy (62% women; mean age=66.93 years). Age groups included were 50 to 60 years in two studies (Fong & Ng, 2006; Tsang et al., 2004); 61 to 70 years in 15 studies (Brooke-

Wavell & Cooling, 2008; Buatois et al., 2007; Dewhurst et al., 2014; Fong et al., 2014; Gao et al., 2011; Gaudagnin et al., 2015; Gyllensten et al., 2010; Tsang & Hui-Chan, 2004; Tsang & Hui-Chan, 2005; Tsang & Hui-Chan, 2006; Tsang & Hui-Chan, 2010; Wayne et al., 2014; Wong et al., 2001; Wong et al., 2011; Zhang et al., 2011); and 71 years or over in eight studies (Aoyagi et al, 2009; Gauchard et al., 1999; Gauchard et al., 2001; Hakim et al., 2004; Hakim et al., 2010; Lu et al., 2013; Perrin et al., 1999; Rahal et al., 2015).

The majority of studies reported anthropometric characteristics such as weight (14 studies, mean weight=60.9, SD=9.4kg), height (14 studies, mean height = 1.57, SD = 1.57m), and body mass index (six studies, mean BMI = 29.5, SD = 3.7), but only two studies reported marital status (Buatois et al., 2007; Hakim et al., 2010); and only one reported ethnicity, race, and education (Wayne et al., 2014).

2.5.2.3 Participant setting

All studies included participants that resided in the community.

2.5.2.4 Physical activity

All PA interventions were land based except for two studies that included mixed PA with a component of swimming (Buatois et al., 2007; Perrin et al., 1999). Sixteen studies included 3D PA (e.g. dance and Tai Chi) (n=842 participants) (Dewhurst et al., 2014; Fong & Ng, 2006; Fong et al., 2014; Gyllensten et al., 2010; Hakim et al., 2010; Hakim et al., 2004; Lu et al., 2013;

Rahal et al., 2015; Tsang et al., 2004; Tsang & Hui-Chan, 2004; Tsang & Hui-Chan, 2005; Tsang & Hui-Chan, 2006; Wayne et al., 2014; Wong et al., 2001; Wong et al., 2011; Zhang et al., 2011).

Ten studies included 'General' PA (e.g. walking, cycling) (n=505 participants) (Aoyagi et al., 2009; Brooke-Wavell & Cooling, 2008; Buatois et al., 2007; Gao et al., 2011; Gauchard et al., 1999; Gauchard et al., 2001; Gauchard et al., 2003; Gaudagnin et al., 2015: Perrin et al., 1999; Tsang & Hui-Chan, 2010).

None of the studies reported adverse events as a result of the PA or adherence to the PA protocol.

2.5.2.5 Physical activity measurement

Only one study used an objective measure of PA, an accelerometer, measuring steps per day (Aoyagi et al., 2009), whilst nine used a variety of validated questionnaire based measures such as the Rapid Assessment of Physical Activity (RAPA), a nine item scale relating to level and intensity of PA (Dewhurst et al., 2014); the Physical Activity Status Score (PASS) a 11-point scale measuring PA duration by a combination of the minutes of exercise per week and the intensity (heavy, modest, or none) (Wayne et al., 2014); the Minnesota Leisure Time Physical Activity Questionnaire (MLTPAQ) measuring the energy expended in PA categorised according to their metabolic equivalent (MET) (Gao et al., 2011; Gyllensten et al., 2010; Tsang et al., 2004; Tsang & Hui-Chan, 2004; Tsang & Hui-Chan, 2005; Tsang & Hui-Chan, 2006; Tsang & Hui-Chan, 2010); and 16 did not specify the tool used (Brooke-Wavell & Cooling, 2008; Buatois et al., 2007; Fong et al, 2006; Fong et al., 2014;

Gaudagnin et al., 2015; Gauchard et al., 1999; Gauchard et al., 2001; Gauchard et al., 2003; Hakim et al., 2004; Hakim et al., 2010; Lu et al., 2013; Perrin et al., 1999; Rahal et al., 2015; Wong et al., 2001; Wong et al., 2011; Zhang et al., 2011).

2.5.2.6 Physical activity duration

All studies included a less active group and a more active group and long-term practice of PA ranging from one to 21 years and over, with two identifying one to five years (Aoyagi et al., 2009; Fong & Ng, 2006); eight identifying six to ten years (Dewhurst et al., 2014; Fong et al., 2014; Hakim et al., 2004; Lu et al., 2013; Tsang et al., 2004; Tsang & Hui-Chan, 2004; Tsang & Hui-Chan, 2006; Zhang et al., 2011); one identifying 11-15 years (Tsang & Hui-Chan, 2010); one identifying 16-20 years (Wong et al., 2001); and one identifying 21 years and over (Buatois et al., 2007). Thirteen studies did not specify length of practice (Brooke-Wavell & Cooling, 2008; Gao et al., 2011; Gaugagnin et al., 2015; Gauchard et al., 1999; Gauchard et al., 2001; Gauchard et al., 2003; Gyllensten et al., 2010; Hakim et al., 2010; Perrin et al., 1999; Rahal et al., 2015; Tsang & Hui-Chan, 2005; Wayne et al., 2014; Wong et al., 2011).

2.5.2.7 Balance

Overall, studies included multiple balance measures, except for three that included only one measure (Buatois et al., 2007; Gaudagnin et al., 2015; Tsang & Hui-Chan, 2004). A detailed description of balance measures is included in Appendix VII.

Primary balance measures

Table 2.3 (page xxxi) shows a summary of the indirect measures of balance reported in observational studies. The following provides a narrative description of the measures:

Sixteen studies included indirect measures relating to the neuromuscular system (n=961 participants) (Aoyagi et al., 2009; Brooke-Wavell & Cooling, 2008; Dewhurst et al., 2014; Fong et al., 2014; Fong & Ng, 2006; Gao et al., 2011; Gauchard et al., 1999; Gaudagnin et al., 2015; Gyllensten et al., 2010; Hakim et al., 2004; Hakim et al., 2010; Tsang et al., 2004; Tsang & Hui-Chan, 2005; Tsang & Hui-Chan, 2010; Wayne et al., 2014; Zhang et al., 2011).

Thirteen studies included indirect measures of cognitive function (n=805 participants) (Brooke-Wavell & Cooling, 2008; Fong et al., 2014; Fong & Ng, 2006; Gao et al., 2011; Gyllensten et al., 2010; Hakim et al., 2010; Hakim et al., 2004; Lu et al., 2013; Tsang et al., 2004; Tsang & Hui-Chan, 2004; Tsang & Hui-Chan, 2005; Wayne et al., 2014; Wong et al., 2011).

Only one observation study (Fong & Ng, 2006) included indirect measures across all the systems required for balance assessment (neuromuscular, cognitive and sensory).

Only three studies included any sensory system measures (n=131 participants) (Fong & Ng, 2006; Gauchard et al., 2001; Tsang & Hui-Chan, 2004) and these included proprioception measures. Only one study (Brooke-Wavell & Colling, 2008) reported fall rate.

Some studies met our inclusion criteria but were excluded from the analyses due to inadequate data and the authors provided no further information on request (n=159 participants) (Gauchard et al., 2003; Rahal et al., 2015; Wong et al., 2001).

Results were estimated from graphical information in seven studies (n=429 participants) (Buotois et al., 2007; Fong et al., 2006; Gaudagnin et al., 2015; Gauchard et al., 1999; Gauchard et al., 2001; Perrin et al., 1999; Wong et al., 2011).

Secondary outcome measures

Table 2.4 (page xxxii) shows a summary of the direct measures of balance reported in observational studies. The following provides a narrative description of the measures:

Three studies used the Sensory Organisational Test (SOT) which measures ability to use visual, vestibular and somatosensory inputs and to suppress inappropriate outputs (described in Appendix VII) (Buatois et al., 2007; Gao et al., 2011; Tsang et al., 2004) (n=139 participants).

Force platforms for the measurement of sway for static or dynamic balance was used in 17 studies (n=1028 participants) (Aoyagi et al., 2009; Brooke-Wavell & Cooling, 2008; Dewhurst et al., 2014; Gao et al., 2011; Gauchard et al., 2001; Gauchard et al., 2003; Gyllensten et al., 2010; Lu et al., 2013; Perrin et al., 1999; Rahal et al., 2015; Tsang & Hui-Chan , 2004; Tsang & Hui-Chan, 2005; Tsang & Hui-Chan, 2006; Tsang & Hui-Chan, 2010; Wayne et al., 2014; Wong et al., 2001; Wong et al., 2011).

The ability to maintain balance whilst standing on a tilt board was measured in two studies (n=113 participants) (Fong & Ng, 2006; Perrin et al., 1999).

2.5.2.8 Quality

Table 2.5 (page xxxiii) and Figure 2.2 (page xliv) present a summary of the risk of bias of included observational studies using the Newcastle Ottawa Scale (NOS). The two independent reviewers agreed on 88.5% of the scores awarded to each study, but the three remaining scores were resolved following discussion. In general, studies were of moderate quality with the majority scoring five to seven stars (n=14 studies or 54%). All studies rated poor on comparability of participants, and the majority (n=14 studies) failed to provide details relating to selection process. The measures of balance included in studies were validated and stated in the main objective across all studies.

2.5.2.9 Selection

Representativeness of the sample

The main method of recruitment used was convenience or volunteering sampling (e.g. local clubs, media, or larger cohort studies) (n=21 studies). Three studies provided no sampling description (Dewhurst et al., 2014; Wayne et al., 2014; Wong et al., 2011), and one study, whilst indicating convenience sampling had been used, offered no description (Zhang et al., 2011). One study used anamnesis by a geriatrician (Rahal et al., 2015).

2.5.2.10 Sample size

One study (Wayne et al., 2014) justified the sample size and one claimed to have carried out a sample size justification but provided no detail (Zhang et al., 2011).

2.5.2.11 Non-respondents

No studies provided data on the number or characteristics of non-responders.

2.5.2.12 Assessment of the exposure

Ten studies identified a validated tool used for the assessment of PA level, with 16 providing no description of the tool used.

2.5.2.13 Comparability

Comparability between groups was adequate in 15 studies (Aoyagi et al., 2009; Brooke-Wavell & Cooling, 2008; Buatois et al., 2007; Fong & Ng, 2006; Fong et al., 2014; Gao et al., 2011; Gyllensten et al., 2010; Lu et al., 2013; Tsang & Hui-Chan, 2005; Tsang & Hui-Chan, 2006; Tsang & Hui-Chan, 2010; Wayne et al., 2014; Wong et al., 2001; Wong et al., 2011; Zhang et al., 2011), but poor in the remaining 11 studies.

2.5.2.13 Outcome assessment

All studies used objective outcome measures of balance which were validated.

2.5.2.14 Statistical testing

Twenty-three studies used appropriate statistical tests and reported all results whilst one study (Dewhurst et al., 2014) only presented one out of three outcomes, and two studies (Rahal et al., 2015; Gauchard et al., 2003) presented only median results for outcomes.

2.5.2.15 Effects of more PA versus less PA

Tables 2.6 and 2.7 show pre- and post-sensitivity analyses results for indirect and direct measures of balance respectively. Post-sensitivity analyses are discussed below.

Primary outcomes

1. More active versus less active groups (indirect measures of balance) (Table 2.6, page xxxiv).

Initial analyses included 16 variables (20 studies; n=1,053 participants). Sensitivity analysis removed five variables (which are excluded from Table 2.6, page xxxiv) due to their high risk of bias (maximal walking speed, functional reach in back, left and right directions, and range of motion), resulting in 11 variables being analysed (13 studies; 733 participants).

Sensitivity analyses showed significant differences between more and less active groups for two variables (preferred walking speed and SLS), which were not identified in initial analyses, but otherwise did not alter findings.

Neuromuscular measures

Table 2.6 (page xxxiv) shows that more active groups achieved faster gait speed (SMD 0.66 m/s; 95% CI 0.26 to 1.06 m/s); better results for two measures of strength using ultra sound tests (SMD 0.57; 95% CI 0.25 to 0.89) and isometric knee extension tests (SMD 0.64; 95% CI 0.35 to 0.94); better results for three measures of functionality with longer time on SLS test (SMD 1.17s; 95% CI 0.74 to 1.60s), higher scores on ABC (SMD 1.47; 95% CI 0.70 to 2.25), and faster time taken to complete the TUG test (SMD -0.70s; 95% CI -1.03 to -0.37s); and better results for one measure of stability with greater distances achieved for the functional reach test (forward) (SMD 0.80m; 95% CI 0.48 to 1.11m).

Sensory measures

Less active groups achieved statistically significant better results for one sensory measure of balance with better results on knee joint repositioning tests (SMD -1.37; 95% CI -2.29 to -0.45). There was no statistically significant difference between more active and less active groups for neuromuscular measures such as handgrip strength or cognitive measures such as MMSE scores or reaction time.

Secondary outcomes

1. More active versus less active groups (direct measures of balance) (Table 2.7, page xxxvi).

Twelve variables were included in analyses (14 studies; n=801 participants) (Table 2.7, page xxxvi: analyses highlighted*). However, for sensitivity analyses three studies were removed, due to high risk of bias (n=162 participants) leaving ten variables (11 studies; n=639 participants) for analysis: significance levels decreased for static body stability eyes open and eyes closed (speed).

More active groups achieved statistically significant better results in three secondary outcome measures, with better tilt board results on directional control (SMD 1.02; 95% CI 0.47 to 1.58%), and maximum excursion (SMD 1.09; 95% CI 0.57 to 1.60%) as well as SOT visual ratios (SMD 0.13; 95% CI 0.03 to 0.22%).

There was no statistically significant difference between more and less active groups for other measures of static or dynamic balance.

Subgroup analysis

There was insufficient similarity among studies or common outcomes to perform subgroup analyses.

2.5.3 Intervention studies

Table 2.8 (page xxxviii) shows a summary of the characteristics of included intervention studies.

2.5.3.1 Design, sample size, and location

Due to the inclusion criteria only four randomised controlled trials (RCTs) were included (Paillard et al., 2004; Wayne et al., 2014; Santos Mendes et al., 2011; Yang et al., 2007). Sample size ranged from 20 (Santos Mendes et al., 2011) to 60 (Wayne et al., 2014) with an average of 38 participants, and only one study (Wayne et al., 2014) justified sample size. Of the four studies, one was US based (Wayne et al., 2014) and the country for the remainder was not specified.

2.5.3.2 Participants

Participants across all studies were defined as healthy and resided in the community (62% women; mean age=68.78 years). Average age of participants was 61-70 years in three studies (Paillard et al., 2004; Santos Mendes et al., 2011; Wayne et al., 2014), and 71 years or over in one study (Yang et al., 2007).

Three studies provided the anthropometric characteristics of height (two studies, mean height= 1.7m), weight (two studies, mean weight = 70.85kg), and/or body mass index (BMI) (one study, mean BMI = 26.5 (SD 5.5), but there was a lack of more detailed demographic information with only one study providing ethnicity, race, and education information (Wayne et al., 2014).

2.5.3.3 Physical activity

All studies included a less active group and a more active group, and all PA interventions were land based where two included '3D PA' (n=109 participants) (Wayne et al., 2014; Yang et al., 2007), and two included 'General PA' (n=41 participants) (Paillard et al., 2004; Santos Mendes et al., 2011).

Only one study used a validated PA assessment tool (PASS) (Wayne et al., 2014).

Intervention duration ranged from a minimum of three months (Paillard et al., 2004; Santos Mendes et al., 2011) to a maximum of six months (Wayne et al., 2014; Yang et al., 2007). All four provided results at baseline and results were also provided at post-trial commencement at three months (Paillard et al., 2004), at four months (Santos Mendes et al., 2011), at both two and six months (Yang et al., 2007), and at both three and six months' (Wayne et al., 2014).

None of the studies reported adverse events because of the PA or adherence to the PA protocol.

2.5.3.4 Balance

Table 2.9 (page xli) provides a summary of the indirect balance measures used across included intervention studies. A detailed description of balance measures is included in Appendix VII.

All studies included a neuromuscular balance measure, but only one included a measure of the cognitive system (MMSE) (Wayne et al., 2014), and none included sensory system measures. One study did not include any indirect measures of balance (Santos Mendes et al., 2010).

Secondary outcome measures

Table 2.10 (page xlii) provides a summary of the direct balance measures used across included intervention studies. One study used the SOT (Yang et al., 2007) (49 participants), and three used force plate platforms (Paillard et al., 2004; Santos Mendes et al., 2011; Wayne et al., 2014).

2.5.3.5 Quality

Figure 2.3 (page xlv) shows a summary of review authors' judgements for each risk of bias item for each study, and Figure 2.4 (page xlvi) shows the distribution of judgements across studies for each risk of bias item. The two independent reviewers agreed on 75% of the scores awarded. The one study on which they disagreed on was resolved following discussion. Overall, the figures suggest a high risk of bias for all studies.

Allocation

Sequence generation

Methods of randomisation included permuted-block randomisation, stratification by sex in intervention group only, randomisation by a computer programme, or were unclear. Risk was judged as low in one study (Wayne et al., 2014), unclear in one (Yang et al., 2007), and high in two (Paillard et al., 2004; Santos Mendes et al., 2011).

Concealment

Allocation concealment was adequate in one study (Yang et al., 2007); unclear in two studies (Paillard et al., 2004; Wayne et al., 2014), and high risk in one study (Santos Mendes et al., 2011).

Blinding

Participants and personnel

It was not possible to blind participants or personnel and, therefore, the risk is high across all studies.

Outcome assessments

Blinding of outcome assessment was judged to be low risk across two studies as objective assessments were used, and blinding was unlikely to have influenced the outcome (Paillard et al., 2004; Yang et al., 2007), unclear in one study (Santos Mendes et al. 2011) and high risk in one study (Paillard et al., 2004).

Incomplete outcome data

Two studies provided adequate information relating to incomplete outcome data (Wayne et al., 2014; Paillard et al., 2004), whilst one study provided no information, so risk was unclear (Santos Mendes et al., 2011) and one study was judged as high risk (Yang et al., 2007).

Selective reporting

Three studies appear to have low risk of selective reporting (Paillard et al., 2004; Santos Mendes et al., 2011; Yang et al., 2007), but one study reported a subset of the outcomes so was judged a high risk of selective reporting (Wayne et al., 2014).

Other potential sources of bias

Publication bias

Due to an insufficient number of studies in each outcome measure the possibility of exploring publication bias in included studies using funnel plots was not performed.

Sample size and power analysis

All studies included 60 or less participants (maximum 60; minimum 21) and only one provided a sample power calculation (Wayne et al., 2014).

Conflict of interest

It was unclear for all studies whether there was a conflict of interest and only one study declared a funding source (Wayne et al., 2014).

2.5.3.6 Effects of more PA versus less PA

Due to the limited number of studies and lack of common outcomes, a best evidence synthesis was explored (Van Tulder et al., 2000).

Key findings relating to direct measures of balance

Two studies reported direct measures (Paillard et al., 2004; Wayne et al., 2014), but only one study provided these measures post-intervention measuring neuromuscular system health using gait speed only (Paillard et al., 2004), and found that walking improved gait speed in more active groups. However, the study was at high risk of bias (Higgins et al., 2011) and low methodological quality (level 3 where only one RCT supports the finding with low or high risk of bias) (Van Tulder et al., 2000) and so provides limited evidence.

Key findings relating to secondary measures of balance

All four studies reported secondary measures of balance (e.g. SOT vestibular, BoS and static and dynamic balance measures), and found that intervention groups had better balance scores. All studies were at high risk of bias (Higgins et al., 2011) and of low methodological quality (Van Tulder et al., 2000), and so evidence is again limited.

Key findings overall

There is limited evidence that free-living PA improves measures of balance in older healthy community-dwelling adults.

Subgroup analyses

The heterogeneity in the nature of the outcome data relating to age, type of PA and duration of effect meant that it was not possible to explore the effects of PA in relation to these variables.

2.6 Discussion

2.6.1 Principal findings

The primary objective of this systematic review was to explore the association between free-living PA and balance in older healthy community-dwelling adults. This review summarises two types of evidence and found that most of the evidence was from cross-sectional studies (26 studies) of moderate methodological quality, and a much smaller proportion was from RCTs (four studies) of low methodological quality. The overall finding was that older adults who engaged in more free-living PA performed better on individual measures of balance.

The evidence from cross-sectional studies showed that free-living PA (Howley, 2001; WHO, 2011; 2012; 2013) is beneficial to balance in older healthy community-dwelling adults (≥50 years), where more active groups experienced better performance on indirect measures of gait speed, strength, functionality and flexibility, and on direct measures of directional control, maximum excursion and SOT visual ratios. These findings extend the findings from previous longitudinal research exploring PA and physical performance by Cooper et al. (2011), that found that habitual leisure-time PA carried out over

(17 years) can improve neuromuscular measures of strength in middle aged adults (36-53 years).

Additionally, limited evidence from a small number of RCTs suggests that free-living PA improves measures of balance in the short-term (three-six months) in older healthy community-dwelling adults which extends the findings from previous research, that short-term (three-six months) exercise, a sub-category of PA, improves balance performance in older unhealthy adults (Gillespie et al., 2012; Howe et al., 2011).

A secondary objective of this review was to identify the characteristics of physical activity associated with balance performance in older healthy community-dwelling adults, and the results showed that the type of PA that improved measures of balance were habitual leisure time activities such as ‘3D PA’ (Lamb et al., 2005) such as tai chi and dance, and ‘General PA’ (Lamb et al., 2005) such as walking and cycling, with a duration of three months to 21 years. Therefore, leisure-time physical activity has both a short-term benefit as well as a cumulative benefit to balance performance. It was not possible to identify the frequency and intensity of free-living physical activity required to benefit balance due to the lack of information reported in included studies, but this review extends our understanding that it is not just exercise that can benefit balance (Gillespie et al., 2012; Howe et al., 2011).

2.6.2 Methodological quality

As previously highlighted, a body of clinical evidence supports that exercise can benefit balance (Gillespie et al., 2012; Howe et al., 2011; Stubbs et al., 2015), but it is evident from this review that few RCTs have explored free-living PA and balance, with most evidence derived from observational studies. Therefore, there is insufficient clinical trial data on which to base clear conclusions.

Furthermore, research suggests that the effects of free-living PA require a longer duration of study than that afforded by RCTs (Morris et al., 1997). For example, this review included studies that explored free-living PA of between three months' and 21 years' duration and in particular, found that 'General PA' such as walking can benefit measures of balance (Aoyagi et al., 2009; Brooke-Wavell & Cooling, 2008; Buatois et al., 2007; Gao et al., 2011; Gauchard et al., 1999; Gauchard et al., 2001; Gauchard et al., 2003; Gaudagnin et al., 2015; Paillard et al., 2004; Perrin et al., 1999; Santos-Mendes et al., 2011; Tsang & Hui-Chan, 2010). In contrast, Howe et al.'s (2011) systematic review of RCTs found no evidence that free-living PA such as walking or cycling, of up to 6 months' duration, improved measures of balance in older unhealthy adults. Thus, the benefits realised from free-living PA may only be realised over time, and future research should consider the appropriateness of the study design involved in exploring associations between free-living PA and balance.

Additionally, the observational studies included in this review were mainly cross-sectional studies (25 of 26 studies) and, therefore, no causal relationship between free-living PA and balance can be determined. Also, participants were either volunteers or recruited using convenience sampling, therefore the generalisability of the findings is limited. Future research exploring the cumulative benefits of free-living PA should include longitudinal studies with a robust design to overcome some of the methodological drawbacks of cross-sectional designs.

2.6.3 Strengths and limitations

A strength of this review is that it considers balance as a multidimensional construct (Horak, 2006; Thomas et al., 2014) rather than a single system, and as a result, includes measures across neuromuscular, cognitive and sensory body systems. Thus, this review measures balance holistically. However, it is evident that whilst this review sought to include measures from multiple body systems, the majority of studies focused on neuromuscular measures (19 of 30 studies) and a smaller proportion included cognitive measures (ten of 30 studies), and even less included sensory measures (three of 30 studies). Overall, only one observational study included measures across all three body systems required for balance.

Additionally, this study found no effect for cognitive measures relating to PA level, but research relating to PA and cognition have provided varying results. For example, Zhu et al.'s (2016) study found that higher levels of objectively measured PA, using an accelerometer, are associated with 36% lower incidence of cognitive impairment, and better maintenance of memory and executive function over three years in older healthy and unhealthy white adults (n=6452 participants; mean age 69.7(8.5) years; 55.3% women; 30.5% black), whilst Young et al.'s (2015) systematic review of 12 RCTs (754 participants; 8-26 week duration) of aerobic exercise showed no cognitive benefits in healthy adults. These mixed results may be due to the inclusion of healthy older adults in the present study and Young et al.'s systematic review. Furthermore, this study found that less active participants had better repositioning results than more active participants. This finding may be because of a lack of protocol description and adherence across studies or due to the method of measurement which has shown to impact on the reliability of the results (Piriyaprasarth et al., 2008). As a result, future studies should seek to include measures across all body systems required for balance and include both healthy and unhealthy populations. Additionally, studies should include a description of the protocol used to assess the outcome and confirm adherence to it to improve the reliability of the results.

Studies in the review reported validated measures for both balance and PA. However, most measures of PA were subjective where different forms of questionnaires were used, except for one study which included an objective measure of PA using an accelerometer (Aoyagi et al., 2009). In contrast, the balance measures included were mainly objective. For example, objective measures of balance such as gait speed (Graham et al., 2008), TUG (Podsialdo & Richardson, 1991; Tinetti et al., 1995), Single legged stance (SLS) (Tinetti et al., 1995), isometric knee extension (Toonstra & Mattacola, 2013), ultrasound (Lord et al., 1994), forward reach test (Newtown, 2001), and knee joint repositioning test (Relph & Herrington, 2015) were included in this review. So, future research should seek to include both validated and objective measures of balance and PA to ensure both components are fully and adequately measured, thus reducing any bias due to self-reporting and or recall in the results (Hillsdon et al., 2005).

There are also some limitations that should be considered. For example, studies provided limited information relating to participant demographics, therefore, limiting investigation of how, for example, socioeconomic status might affect PA level and balance. Additionally, results were estimated from graphical information in seven observational studies (n=429 participants) which may give rise to bias in the results. Furthermore, the sample size for both cross-sectional studies and RCTs were small ranging from 20-170 participants, and only one study carried out *a priori* sample size calculation to inform the study population size (Wayne et al., 2014). This may give rise to Type II errors where the null hypothesis is wrongly accepted. It is therefore important that future studies should include demographics, include only reported data rather than graphical representation of data, and perform *a priori* sample size calculation to conduct adequately powered studies (Brewer & Sindelar, 1988) to improve the results and widen the scope of research.

Additionally, there is no consensus relating to the most suitable balance measures as earlier indicated in Chapter one. This review included multiple balance measures across different body systems to comprehensively measure balance (n=40 different balance measures) (Horak, 2006), and as confirmed by other studies (Howe et al., 2011), the number of different outcome measures restricted the ability to compare and pool results. Therefore, future research in this emerging area should consider establishing a consensus of relevant balance measures across all body systems to aid analysis and fully understand the effects of free-living PA on balance.

2.6.4 Future research and clinical implications

In summary, this review suggests that free-living PA improves balance performance in older healthy community-dwelling adults both in the short-term and long-term. Further studies need to consider the intensity and frequency of physical activity required to benefit balance using objective measures. Additionally, further research that incorporates higher quality studies is warranted, and the inclusion of longitudinal studies that provide large samples of participants using robust selection processes, and appropriate data over multiple time points, should be considered. For example, studies such as NICOLA (https://www.mrc.ac.uk/research/facilities-and-resources-for-researchers/cohort-directory/northern-ireland-cohort-longitudinal-study-of-ageing-nicola/), TILDA (http://tilda.tcd.ie/), and ELSA (https://www.elsa-project.ac.uk/) include large samples of community-dwelling participants (≥50 years) (8,500, 8,504 and 11, 391 respectively); provide data across multiple timepoints (between three to 11 years); adhere to the Gateway to Global Ageing Initiative (https://g2aging.org/) which improves the harmonisation of balance outcomes, therefore reducing the variability of outcomes and improving comparability of results; and include balance measures across multiple body systems that are objective and validated.

2.7 Conclusion

In conclusion, there is limited evidence from low to moderate quality studies that free-living PA improves measures of balance in older community-dwelling healthy adults, particularly in respect of fall prevention. Additionally, research suggests that free-living PA may require a longer timeframe over which the benefits can be realised than that generally afforded by for example, clinical trials. Furthermore, studies assessing the effects of PA on balance do not measure balance comprehensively across those body systems required for balance (as identified in Chapter one, section 1.3.2). Therefore, robustly designed longitudinal studies that explore the effects of PA on a comprehensive measurement of balance, are warranted to improve the overall robustness of the findings. Therefore, the study carried out in Chapter three investigates the effects of free-living PA in older community-dwelling adults using multiple indirect measures of balance across neuromuscular, cognitive, and sensory body systems and PA measured using the International Physical Activity Questionnaire (IPAQ). It uses longitudinal data (over a two-year period from 8,504 participants) from TILDA to develop a predictive model of PA and balance in older adults (≥50 years).

Table 2.1
Excluded studies summary from full text review (n=52 studies)

Author	No balance measure	No PA measure	Age <50yrs	Structured exercise/lab based	Unhealthy population	No comparison group	Excluded Study type	Other
Akosile et al., 2014	X							
Albert et al., 2014		X						
Alexander et al., 2008							Discussion paper	
Alpert et al., 2009						X		
Aranda-Garcia et al., 2015						X		
Bauman et al., (2000)							Discussion paper	
Brouwer et al., (2004)		X						
Busing, 2005				X				
Cancela et al., (2002)								Non-english
Dattilo et al., (2012)					History of falls			
De Rekeneire et al., (2003)		X						
Demura et al., (2012)		X						
Domaradzki et al., (2014)		X						
Earles et al., (2001)					Arthritis			
Egerton et al., (2009)		X						
Ekman et al., (2001)								Not available
El Haber et al., (2008)					History of falls			
Faude et al., (2012)					History of falls			
Frandin et al., (1995)		X						
Freitas et al., (2013)		X						
Gill et al., (2016)					Disabilities			
Goutier et al., (2010)		X						

Table 2.1(continued)
Excluded studies summary from full text review (n=52 studies)

Author	No balance measure	No PA measure	Age <50yrs	Structured exercise/lab based	Unhealthy population	No comparison group	Excluded Study type	Other
Graafmans et al., (2003)		X						
Gressl et al., (2007)	X							
Guan et al., (2011)			X					
Gudlaugsson et al., (2013)				Walking performed indoors & includes endurance training				
Gustafson et al., (2000)					Hospitalised population			
Halvarsson et al., (2011)				Structured exercise programme				
Halvarsson et al., (2013)					History of falls			
Halvarsson et al., (2011)					History of falls			
Heesch et al., (2008)					History of falls			
Karinkanta et al., (2005)					Hospitalised population			
Kelsey et al., (2010)					History of falls			
Kermode-Scott, (2002)							Discussion paper	
Kim et al., (2014)						X		
Kolt et al., (2011)							Protocol	
Kramer et al., (2014)				Flexibility training				

Table 2.1(continued)
Excluded studies summary from full text review (n=52 studies)

Author	No balance measure	No PA measure	Age <50 yrs	Structured exercise/lab based	Unhealthy population	No comparison group	Excluded Study type	Other
Melo et al., (2011)				Walking is part of the exercise programme but includes strengthening training exercise				
Melzer et al., (2003)							Retrospective case control study	
Melzer et al., (2013)							Retrospective case control study	
Mendoza-Ruvalcaba & Arias-Merino (2015)				Strength and balance exercise				
Meuleman et al., (1992)					X			
Montero-Odasso et al., 2005		X						
Musselman et al (2005)		X						
Paterson et al (2011)		X						
Pober et al (2002)						X		
Rosano et al (2005)						X		
Tsang et al (2011)								No UK locations
Valentine et al (2009)						X		
Van Dijk et al., 2013						X		
Wayne et al., (2015)								Duplicate
Yamada et al (2009)						X		

Table 2.2
Study characteristics of included observational studies (n=26 studies)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	**More active (MA) versus less active (LA)**	**Measure, Duration, Intensity**	**Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL); Stability (ST)**	**Cognitive**	**Sensory**	**Other**	
Aoyagi et al., 2009	-Prospective cohort – 1yr -Recruitment: Nakanojo study -Conflict of Interest: N/k -Source of funding: declared	N: 170 -Age: 72.6±4.6yrs (65-84yrs) -55% women -Height(m): 1.53±0.08 -Weight(kg): 54.3±8.6 -BMI: 23.3±3.3 -Japan -Community setting -Written informed consent	MA group: 65-74yr group LA group: 75-84yr group	Accelerometer MA group: 7,190±2,491 steps per day LA group: 5,482±2,829 steps per day	**Indirect measure** **(G)** Walking speed (preferred & maximal) (5m) (velocity - m/s) **(S)** Handgrip test (dynamometer) (force - n) **(S)** Isometric knee extension (dynamometer) (torque – N*m/kg) **(ST)** Functional reach test (distance - m) **Direct measure** *Static balance test:* Total body stability (eyes open/closed) (sway distance - m)	n/a	n/a	n/a	Physical fitness except handgrip and total body sway were greater for MA group (65-74yr).

Table 2.2(continued)
Study characteristics of included observational studies (n=26 studies)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	**More active (MA) versus less active (LA)**	**Measure, Duration, Intensity**	**Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)**	**Cognitive**	**Sensory**	**Other**	
Brooke-Wavell & Cooling, 2008	-Cross sectional -Recruitment: local bowls clubs; media & friends & family) -Conflict of interest: N/k -Source of funding: N/k	-N: 74 -Age: 68.3±4.65yrs (60-75yrs) -100% women -Weight(kg): 69.2±10.1 -BMI (kg/m²): 26.95±3.9 -Community setting -Written informed consent	MA group: Bowlers LA group: non-bowlers	MA group: 2-3+ hours of PA per week LA group: less than 3 hours PA per week	**Indirect measure** **(S)** Isometric knee & hip extension (scat & force meter) (force - n) **(S)** BUA of the calcaneus (Osteometer) (dB/MHz) **(FU)** TUG (3m) (time - s) **(FL)** Range of Motion: shoulder & ankle (goniometer) (degrees˚) **Direct measure** *Static balance test:* Total body stability (eyes open/closed (distance - mm)	Reaction time (s)	n/a	Falls	MA group had better postural stability, muscle strength and flexibility

Table 2.2(continued)
Study characteristics of included observational studies (n=26 studies)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Buatois et al., 2007	-Cross sectional -Recruitment: cohort from a larger study -Conflict of Interest: N/k -Source of funding: declared	-N: 130 -Age: 70.3±4.3yrs -41% women -BMI (kg/m²): 26.28±3.75 -Community setting -Written informed consent	MA group: PA -walking, cycling, swimming, gymnastics PA experience: 28±9.5yrs LA group: no PA	MA group: 1-2 hours per week LA group: no PA	**Direct measure** Sensory Organisation Test (equilibrium scores and composite score)	n/a	n/a	n/a	Sensory conflicting conditions were more challenging for LA group who swayed more and frequently lost balance than MA group.

Table 2.2(continued)
Study characteristics of included observational studies (n=26 studies)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Dewhurst et al., 2014	-Cross sectional -Recruitment: n/k -Conflict of interest: N/k -Source of funding: N/k	-N: 60 -Mean age: 69.36±2.9yrs (60-80yrs) -100% women -Height(m): 1.58±0.07 -Weight(kg) 64.05±8.15 -BMI (kg/m²):25.95±3.9 Waist(cm)82.45±9.08 Hip (cm): 102.6±7.62 Waist/hip ratio: 0.80±0.2 -Scotland -Community setting -Written informed consent	MA group: Dancers LA group: Non-dancers	RAPA MA group: 2.5hrs hours of PA per week 10yrs Scottish dance experience LA group: 2.5 hours PA per week (no dancing)	**Indirect measure** **(G)** Walking speed (preferred/maximum) (6m) (speed - s) **(FU)** Timed Up & Go (2.44m) (time to complete - s) **(FL)** Range of motion: Chair sit & reach test (distance - cm) **(FL)** Range of motion: Back scratch test (left/right shoulder) (distance - cm) **Direct measure** *Static balance test* Total body stability (sway area -cm²)	n/a	n/a	n/a	No differences in measures of flexibility between groups. Better results for MA group on measures of TUG, walking and sway.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Fong & Ng, 2006	-Cross sectional -Recruitment: N/k -Conflict of Interest: N/k -Source of funding: N/k	-N: 48 -Age: 55.4±11.5yrs -50% women -Community setting -Written informed consent	MA group: tai chi LA group: no tai chi	MA group: 3-6hrs per week 1-3yrs tai chi experience LA group: no tai chi	**Indirect measure** **(FL)** Knee repositioning (electrogoniometer) (˚) (degrees; absolute error) **Direct measure** Tilt board (balance time - s)	Reaction time (electromyography) (ms)	Knee angle repositioning	n/a	MA group had better reaction times, knee joint positioning, and dynamic standing balance than LA group.
Fong et al., 2014	-Cross sectional -Recruitment: martial arts/elderly centres -Conflict of interest: N/k -Source of funding: N/k	-N: 84 -Age: 64.39±11.9yrs -44% women -Weight (Kg): 63.2±11.8 -Height(m): 1.60±0.09 -BMI (kg/m²): 49.3±3.65 -Community setting -Written informed consent	MA group: martial arts LA group: no martial arts	MA group: 2 hours per week of martial arts Experience: 8±9.9yrs LA group: no martial arts	**Indirect measure** **(FU)** Five times sit to stand (time to complete s) **(FU)** Berg Balance Scale (14 items) (max score 56) **Direct measure** Bone ultrasound: arm (SOS T & Z scores)	ABC (16 items)	n/a	n/a	MA had better bone strength, lower limb muscular strength and better functional balance than LA group.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Gao et al., 2011	-Cross sectional -Recruitment: local golf clubs/ community centres -No conflict of interest -Source of funding: declared	-N: 23 -Age: 68.75±6.7yrs (60-80yrs) -0% women -Height(m): 1.6±0.06 China -Community setting -Written informed consent	MA group: Golfers LA group: Non-golfers	MLTPAQ MA group: Light =6; Mod. =4; Heavy =1 LA group: Light =10; Mod. = 2; Heavy =0	**Indirect measure** **(ST)** Functional reach test (forward) (functional reach normalised with body height - %) **Direct measure** Sensory Organisation Test (somatosensory, visual and vestibular ratios)	MMSE (30 items) ABC (mod.) (16 items)	n/a	n/a	MA group had better balance control, reach, postural control, visual & vestibular inputs. No significant difference between somatosensory ratios between groups.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL))	Cognitive	Sensory	Other	
Gauchard et al., 1999	-Cross sectional -Recruitment: existing ageing study -Conflict of interest: N/k -Source of funding: N/k	-N: 40 -Age: 72.7±6.5yrs -70% women -Community setting -Informed consent	MA group: yoga & soft gymnastics LA group: walking	MA group: 90mins per week LA group: 5km per week	**Indirect measure** **(S)** Knee & ankle extension/flexion, dynamometer (power - Nm/s; strength - Nm) **Direct measure** *Dynamic balance test* AP stability (eyes open/closed) (foot displacement - FFT; strategy type - Type 1, 2, & 3)	n/a	n/a	n/a	Regular PA improves measures of strength and postural control.
Gauchard et al., 2001	Cross sectional Recruitment: existing ageing study Conflict of interest: N/k Source of funding: N/k	N: 36 Age: 72.9±6.5yrs 72% women Community setting Informed consent	MA group: yoga & soft gymnastics LA group: walking	MA group: 90mins & 5km walking per week LA group: 5 km per week	**Direct measure** *Static balance test* AP (eyes open/closed) (ratio) *Dynamic balance test* AP stability (eyes open/closed) (component velocities of nystagmus -left, right, total R-MSCV; L-MSCV; T-MSCV; strategy type 1, 2, 3)	n/a	Vestibular tests; caloric/rotational;vestibular reflectivity	n/a	Inactivity causes poor balance, vestibular hypo excitability and dependency on visual afferent. PA such as yoga improves dynamic postural control.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	**More active (MA) versus less active (LA)**	**Measure, Duration, Intensity**	**Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)**	**Cognitive**	**Sensory**	**Other**	
Gauchard et al., 2003	-Cross sectional -Recruitment: cohort study -Conflict of interest: N/k -Source of funding: N/k	-N: 44 -Median age:73.33yrs (63-85yrs) -100% women -Community setting -Written informed consent	MA group: yoga & soft gymnastics LA group: no PA: walking	MA group: 90 mins per week LA group: n/k	**Direct measures** *Static balance test* Total body stability (sway distance - m; sway area -cm²) AP & ML stability (eyes open/closed) (sway distance - m; sway area - cm˚; ratio - EO/EC)	n/a	n/a	n/a	Regular PA & Proprioceptive PA like yoga is more successful in improving static balance.
Gaudagnin et al., 2015	-Cross sectional -Recruitment: N/k -No conflict of interest -Source of funding declared	-N: 24 -Age: 67.5±5.5yrs 100% women -Height(m): 1.54±0.06; -Weight (Kg): 65.5±10.5 -Brazil -Community setting -Written informed consent	MA group: PA LA group: no regular PA	MA group: at least 150mins per week LA group: no PA	**Indirect measure** **(G)** Walking speed (preferred) (8m) (velocity - m/s)	n/a	n/a	n/a	Active lifestyle improves gait speed.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	**More active (MA) versus less active (LA)**	**Measure, Duration, Intensity**	**Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)**	**Cognitive**	**Sensory**	**Other**	
Gyllensten et al., 2010	-Cross-sectional -Recruitment: community centres -No conflict of interest -Source of funding: N/k	-N: 44 -Age: 69.9±6.85yrs -82% women -Weight(k) 54.8±6.95 -Height(m): 1.55±6.95 -Hong Kong, China -Community setting -Written informed consent	MA group: Tai chi LA group: Non-tai chi	MLTPAQ MA group: Light =4; Mod. =17; Heavy =3 LA group: Light =7; Mod. =12; Heavy =1	**Indirect measure** **(FU)** Body Awareness Scale- Healthy (BAS-H) (25 items) **(FU)** Single Leg Jump Test (yes/no; s) **Direct measure** *Dynamic balance test* Limits of Stability (velocity - °/sec; endpoint/ maximum excursion; directional control - %)	MMSE (mod.) (30 items)	n/a	n/a	MA group - stability limits, single leg stance, stability on landing on one leg.
Hakim et al., 2004	-Cross-sectional -Recruitment: local tai chi clubs/senior centres -Conflict of Interest: N/k -Source of funding: N/K	N: 94 Age: 75.2±7.5yrs (60-96yrs) 84% women Pennsylvania; US Community setting Written informed consent	MA group: Tai chi LA group: No exercise	MA group: 62.5% walk regularly, all take tai chi 1 or more times per week (tai chi experience: 5.6yrs); LA group: no tai chi or walking	**Indirect measure** **(FU)** Timed Up & Go (3m) (time to complete - s) **(FU)** Chair stand test (30s) (number of full stands) **(FL)** Multidirectional reach test (distance - inches)	ABC (16 items)	n/a	n/a	MA group have better balance, confidence, and reach.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Hakim et al., 2010	-Cross- sectional -Recruitment: local tai chi/senior centres -No conflict of interest -Source of funding: N/k	-N: 52 -Age: 74.46±5.09yrs -87% women -Marital status: Single=17%; Married=30%; Divorced=11% Widowed=42% -Community setting -Informed consent	MA group: Tai chi LA group: No exercise	MA: 11.66±5.15 (days/month) LA group: 10.73±9.52 (days/month)	**Indirect measure** **(FU)** Fullerton Advanced Balance Scale (FAB) (10 items) **(FU)** Time Floor Transfer test (time to complete - s) **(FU)** Single leg stance (30s) (balance time - s) **(FL)** Multidirectional reach test (distance - inches)	ABC (16 items)	n/a	n/a	MA group have better scores on FAB and reach test. No differences found on ABC, single leg stance, and timed floor transfer test between groups.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Lu et al., 2013	-Cross sectional -Recruitment: local tai chi clubs/ elderly centres -No conflict of interest -Source of funding declared	-N: 58 -Age: 73.5±5.15yrs -72% women -Height(m): 1.54±0.80 -Weight(kg): 56.95±9.1 -Hong Kong, -China -Community setting -Written consent	MA group: Tai chi LA group: Non-tai chi	MA group: Light=4; Mod. =23; Heavy=1 Minimum of 1.5hrs per week tai chi Tai chi experience: 6.7±4.6yrs LA group: No tai chi: Light=5; Mod. =25; Heavy=0	**Direct measures** *Static balance test* Total body sway (dual and single task) (sway distance - mm; sway area - cm²)	MMSE (30 items) Auditory Stroop test (reaction time (s); error rate (%)		n/a	MA group performed better in stepping down and Stroop tests.
Perrin et al., 1999	-Cross sectional -Recruitment: ageing study -No conflict of interest -Source of funding: N/k	-N: 65 -Age: 71.8s±0.8yrs -66% women -France -Community setting	MA group: walking/ swimming /cycling/ tennis LA group: no PA	MA group: n/k LA group: no PA	**Direct measure** *Static balance test:* Total body stability (eyes open/closed) (sway velocity - cm/s; sway area - cm²) AP/ML stability (eyes open/closed) (sway velocity -cm/s; sway area - cm²) *Dynamic balance test:* Tilt board (Short, medium, and long latency responses)	n/a	n/a	n/a	Balance in EO or EC conditions is better in MA group.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Rahal et al., 2015	-Cross sectional -Recruitment: geriatrician by anamnesis -Conflict of interest: N/k -Source of funding: N/k	-N: 76 -Age: 73.55yrs (60-80yrs) -74% women -Community setting -Written informed consent	MA group: Tai chi group LA group: Dance group	Measure: n/k MA group: up to 3 hrs tai chi per week LA group: up to 3 hrs dance per week	**Direct measure** *Static balance test:* Modified Clinical Test of Sensory Interaction on Balance (mCTSIB) (sway velocity - ˚/s) Unilateral stance (sway velocity - ˚/s) *Dynamic balance test:* Walk across test: (sway speed - cm/s; step width - cm; sway velocity - ˚/s) Sit to stand test: (sway velocity - - ˚/s; weight transfer	n/a	n/a	n/a	MA group had reduced postural sway and thus improved static and dynamic balance.
Tsang & Hui-Chan, 2004	-Cross sectional -Recruitment: tai chi clubs -No conflict of interest -Source of funding declared	-N: 47 -Age: 69.03±6.37yrs 0% women -Height(m): 1.61±6.45 -Weight(kg): 62.65±7.75 -Community setting -Written consent	MA group: Tai chi group experience: 8.4yrs LA group: No exercise	MLTPAQ MA group: Light =7; Mod. =4; Heavy = 1 (PA - Up to 1.5hrs p/w) LA group: Light = 10; Mod. =2; Heavy =0, Walked daily)	**Direct measure** *Dynamic balance test* Limits of stability test (reaction time (s); maximum excursion (%); directional control (%))	MMSE (30 items)	Passive knee joint repositioning test (angle error - ˚)	n/a	MA group had better knee joint position.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	**More active (MA) versus less active (LA)**	**Measure, Duration, Intensity**	**Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)**	**Cognitive**	**Sensory**	**Other**	
Tsang & Hui-Chan, 2005	Cross sectional Recruitment: tai chi clubs/ community centres No conflict of interest Source of funding declared	N: 48 Age: 70.45±5.55yrs 50% women Height(m): 1.55±0.07 Weight(kg): 58.1±9.05 Community setting Written informed consent	MA group: Tai chi LA group: No tai chi	MLTPAQ MA group: Light =17; Mod. =5; Heavy = 2 PA Up to 1.5hrs per week LA group: Light =21; Mod. =3; Heavy =0 (Walked daily)	**Indirect measure** **(S)** Isokinetic knee strength test (dynamometer) (peak torque to body weight ratio) **Direct measure** *Static balance test* AP & ML body stability (body sway angle ˚) *Dynamic balance test* AP & ML body stability (body sway angle (˚))	ABC (16 items)	n/a	n/a	MA group showed better knee muscle strength, less body sway in static standing and perturbed single leg stance and greater balance confidence.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Tsang & Hui-Chan, 2006	-Cross sectional -Recruitment: tai chi clubs/ community centres -Conflict of interest: N/k -Source of funding: N/k	-N: 48 -Age: 70.45±5.55yrs -50% women -Height(m): 1.55±0.09 -Weight(kg): 58.1±17.5 -Community setting -Written consent	MA group: tai chi group Tai chi experience: mean 8.5yrs LA group: No tai chi group	MLTPAQ MA group: Light =17; Mod. =5; Heavy =2 (PA Up to 1.5hrs per week) LA group: Light =21; Mod. =3; Heavy =0 (Walked daily)	**Direct measure** *Static balance test:* Total body stability pre-& post vestibular stimulation (eyes open/closed) (sway distance - cm) AP & ML stability pre-& post vestibular stimulation (eyes open/closed) (velocity -cm/s; amplitude˚)	n/a	n/a	n/a	MA group have better control of body sway along AP direction.
Tsang & Hui-Chan, 2010	-Cross sectional -Recruitment: golf clubs/communi ty centres -No conflict of interest -Source of funding declared	-N: 23 -Age: 68.75±6.7yrs -0% women -Height(m): 1.62±6.95 -Weight(kg): 64.05±8.15 -Community setting -Written consent	Ma group: Golfers Golf experience: 15.2yrs LA group: Non-golfers	MLTPAQ MA group: Light =6; Mod. =4; Heavy =1 (PA Up to 1.5hrs per week) LA group: Light =10; Mod. =2; Heavy =0 (Walked daily)	**Indirect measure** **(FU)** Single leg stance (balance time -s) **(FL)** forward lunge test (average distance of lunge as % of height) **Direct measure** *Dynamic balance test* AP body stability (body sway angle ˚)	n/a	n/a	n/a	MA group had better single/pertu rbed leg stance, smaller sway, larger lunge distance on both legs.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Tsang et al., 2004	-Cross sectional -Recruitment: centres for elderly -Conflict of interest: N/k -Funding: N/k	-N: 60 -Age: 53.33±3.73yrs -50% women -Height(m): 1.57±0.09 -Weight(kg): 58.7±9.7 -Hong Kong, China -Community setting -Informed consent	MA group: Tai chi group Tai chi experience: 7.2yrs LA group: No tai chi group	MLTPAQ MA group: Light =1; Mod. =15; Heavy =4 (PA Up to 3hrs per week) LA group: Light =0; Mod. =15; Heavy = 5 (Walked daily)	**Indirect measure** **(S)** Handgrip test (dynamometer) (strength (Kg)) **Direct measure** Sensory Organisation Test (somatosensory, visual, vestibular ratios)	MMSE (mod.) (30 items)	n/a	n/a	MA group had better postural control in all sensory conditions.
Wayne et al., 2014	-Cross sectional -Recruitment: N/k -Conflict of interest: N/k -Funding: N/k	-N: 87 -Age: 63.48±7.63yrs (50-79yrs) -66% women -White 86%; Non-Hispanic 98% -Education: 18±3.3yrs -BMI (kg/m²) 25±3.9 -Boston, US -Community setting	MA group: Tai chi expert LA group: Tai chi naïve	PASS MA group: 6.0±2.0 (intensity/mins pw) LA group: 4.4±2.2 (intensity/mins pw)	**Indirect measure** **(FU)** Timed Up & Go (s) **(FU)** Single leg stance (s) **Direct measure** *Static balance test:* Total body stability (EO/EC) (sway velocity (mm/s); sway area (mm²)) Dynamic balance test: AP & ML stability (EO/EC sway velocity (mm/s)	MMSE (30 items)	n/a	n/a	MA better sway, single leg stance and TUG.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	**More active (MA) versus less active (LA)**	**Measure, Duration, Intensity**	**Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)**	**Cognitive**	**Sensory**	**Other**	
Wong et al., 2001	-Cross sectional -Recruitment: tai chi club group -Conflict of interest: N/k -Source of funding declared	-N: 39 -Age: 68.47±5.53yrs -69% women -Weight(kg): 64.73±8.03 -Height(m): 1.57±0.08 -Taiwan -Community setting -Informed consent	MA group: tai chi LA group: no tai chi	MA group: tai chi Experience: 15.6±10.5yrs LA group: no tai chi	**Direct measure** *Static balance test* Total body stability (eyes open/closed) (max stability - %; sway velocity - ˚/s) *Dynamic balance test* Total body stability (eyes open/closed) (max stability - %; sway velocity - ˚/s)	n/a	n/a	n/a	MA group had better postural control than LA group.
Wong et al., 2011	-Cross sectional -Recruitment: local tai chi clubs -Conflict of interest: none -Source of funding declared	-N: 86 -Age: 66.93±5.63yrs -62% women -Weight (Kg): 58.65±8 -Height(m): 1.57±0.07 -Taiwan -Community setting -Written consent	MA group: tai chi LA group: no PA	MA group: 162mins per week LA group: no PA	**Direct measure** *Static balance test* Total body stability (eyes open/closed) (max stability - %; sway velocity - ˚/s; ankle strategy - %) *Dynamic balance test* Total body stability (eyes open/closed) (max stability - %; sway velocity - ˚/s; ankle strategy - %)	Reaction time (eye/hand speed - ms)	n/a	n/a	MA group showed better stability, smaller COP velocity, & greater use of ankle strategy.

Table 2.2
Study characteristics of included observational studies (n=26 studies) (continued)

Study Author	Study Design	Study Population	Physical Activity measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Zhang et al., 2011	Cross sectional Recruitment: local tai chi/ walking groups Conflict of interest: nonet Source of funding declared	N: 30 Age: 65.7±4.9yrs 100% women Community setting Written informed consent	MA group: Tai chi group LA group: Walking group	MA group: 7hrs per week of tai chi 8.2yrs tai chi experience LA group: 7hrs per week of walking 8.8yrs walking experience	**Indirect measure** **(FU)** Single leg stance (time spent on one leg during walking (s)) **(G)** Walking speed (preferred) (velocity (m/s)	n/a	n/a	n/a	MA group have better movemen t control, but LA group have better results on single leg stance measures .

Table 2.3
Summary of indirect balance outcomes reported in observational studies

Body system	**Neuromuscular**																	**Cognitive**			**Sensory**		**Other**
Measure	Gait		Strength			Functionality									Stability	Flexibility		Cognition			Proprioception	Vestibular	
Indirect Measures	Preferred walk speed	Max. walk speed	Handgrip test	Isometric Knee Extension	Ultra sound	Chair stand test	Single leg jump test	Five times sit to stand	TUG	FAB	TFT	BAS-H	Single leg stance	Berg Balance Scale	Functional reach(forward/back ward/left/right)	ROM	Forward lunge	MMSE	Reaction Time	ABC	Knee joint repositioning	Vestibular tests	Falls
Aoyagi (2009)	X	X	X	X											X								
Brooke-Wavell (2008)				X	X				X							X			X				X
Dewhurst (2014)	X	X							X							X							
Fong (2006)					X													X			X		
Fong (2014)					X			X						X						X			
Gao (2011)															X			X		X			
Gauchard (1999)				X																			
Gauchard (2001)																						X	
Gaudagnin (2015)	X																						
Hakim (2010)										X	X		X		X					X			
Hakim (2004)						X			X						X					X			
Wayne (2014)									X				X					X					
Gyllensten (2010)							X					X						X					
Lu (2013)																		X	X				
Tsang (2005)				X																X			
Tsang (2004)			X															X					
Tsang (2004)																		X			X		
Tsang (2010)													X				X						
Wong (2011)																			X				
Zhang (2011)	X												X										

Table 2.4
Summary of the direct balance outcomes reported in observational studies

Direct measures	**Force platform & sway indicators (static**	Total body stability (distance)	Total body stability (area)	Total body stability (velocity)	Total body stability (Romberg Quotient)	AP & ML stability (distance)	AP & ML stability (angle)	AP & ML stability (area)	AP & ML stability (force)	**Force platform & sway indicators (dynamic)**	Total body stability (Romberg Quotient)	Total body stability (velocity)	AP & ML stability (backward/forward) (angle)	AP & ML stability (force)	Stability - strategy type	Limits of stability	AP or ML stability only	**Other**	SOT (composite or ratios)	Tilt board
Aoyagi (2009)	**X**	X																		
Brooke-Wavell, (2008)	**X**	X																		
Buatois (2007)																			X	
Dewhurst (2014)	**X**		X																	
Gao (2011)	**X**					X													X	
Gauchard (1999)																				
Gauchard (2001)	**X**				X					**X**	X				X					
Gauchard (2003)	**X**	X	X			X		X												
Fong (2006)																				X
Perrin (1999)	**X**		X	X				X	X											X
Gyllensten (2010)										**X**		X				X				
Lu (2013)	**X**	X	X																	
Rahal (2015)	**X**			X																
Tsang (2005)	**X**						X			**X**			X							
Tsang (2006)	**X**	X					X		X											
Tsang (2004)																			X	
Tsang (2004)										**X**						X				
Tsang (2010)										**X**			X				X			
Wayne (2014)	**X**		X	X						**X**				X						
Wong (2011)	**X**			X						**X**		X								
Wong (2011)	**X**			X																

Table 2.5
Summary of the risk of bias in observational studies using the Newcastle-Ottawa Scale

Study	**Selection** (max. 5 stars)	**Comparability** (max. 2 stars)	**Outcome** (max. 3 stars)	**Total** (max. 10 stars)
Aoyagi et al., 2009	***	*	***	7
Brooke-Wavell & Cooling, 2008	*	*	***	5
Buatois et al., 2007	*	*	***	5
Dewhurst et al., 2014	**		**	4
Fong & Ng, 2006	*	*	***	5
Fong et al.,2014	*	*	***	5
Gao et al., 2011	***	*	***	7
Gauchard et al., 1999	*		***	4
Gauchard et al., 2001	*		***	4
Gauchard et al., 2003	*		**	3
Gaudagnin et al., 2015	*		***	4
Gyllensten et al., 2010	***	*	***	7
Hakim et al., 2004	*		***	4
Hakim et al., 2010	*		***	4
Lu et al., 2013	*	*	***	5
Perrin et al., 1999	*		***	4
Rahal et al., 2015			**	2
Tsang & Hui-Chan, 2004	***		***	6
Tsang & Hui-Chan, 2005	***	*	***	7
Tsang et al., 2004	***		***	6
Tsang & Hui-Chan, 2006	***	*	***	7
Tsang & Hui-Chan, 2010	***	*	***	7
Wayne et al., 2014	***	*	***	7
Wong et al., 2001	*	*	***	4
Wong et al., 2011		*	***	4
Zhang et al., 2011		*	***	4

Table 2.6
Pre- and post- sensitivity analysis results for observational studies (Indirect measures of balance)

Comparison or subgroup	No. of studies	N	Effect size (95% CI)	Heterogeneity
Neuromuscular measure of gait				
*1 Preferred walking speed (m/s).	4	284	0.24 (-0.69, 1.17)	91%
1.1 Preferred walking speed (m/s).	2	194	0.66 (0.26, 1.06)	20%
Neuromuscular measures of strength				
*2 Handgrip (Kg). ++	2	210	1.73 (-1.20, 4.66)	23%
*3 Isometric knee extension.	4	320	0.63 (0.40, 0.87)	35%
3.1 Isometric knee extension.	3	292	0.64 (0.35, 0.94)	25%
*4 Ultrasound.	2	158	0.57 (0.25, 0.89)	0%
Neuromuscular measures of functionality				
*5 Timed Up & Go. (s) Low value indicates better balance.	4	286	-0.76 (-1.01, -0.51)	0%
5.1 Timed Up & Go. (s) Low value indicates better balance.	2	161	-0.70 (-1.03, -0.37)	0%
*6 Single Leg Stance. (s)	4	181	-0.25 (-1.86, 1.37)	95%
6.1 Single Leg Stance. (s)	2	110	1.17 (0.74, 1.60)	0%
*7 Activities of Balance Confidence.	4	220	1.33 (0.73, 1.94)	74%
7.1 Activities of Balance Confidence.	3	155	1.47 (0.70, 2.25)	70%
Neuromuscular measures of stability				
*8 Functional reach (forward) (m).	4	304	1.18 (0.61, 1.75)	74%
8.1 Functional reach (forward) (m).	2	193	0.80 (0.48, 1.11)	0%

Table 2.6 (continued)
Pre- and post- sensitivity analysis results for observational studies (Indirect measures of balance)

Comparison or subgroup	No. of studies	N	Effect size (95% CI)	Heterogeneity
Sensory measures				
*9 Knee joint repositioning (degrees).	2	58	-1.37 (-2.29, -0.45)	59%
Cognitive measures				
*10 Mini Mental State Exam. ++	4	229	0.37 (-0.35, 1.09)	60%
*11 Reaction time (s). Low value indicates better balance.	3	198	-0.75 (-1.45, -0.04)	83%
11.1 Reaction time (s). Low value indicates better balance.	2	132	-0.41(-0.84, 0.01)	33%

Note:

Data are shown in the table for 11 variables. For some variables there were two sets of data, the first set of data identified with * includes all available data, whereas the second set of data excludes studies at high risk of bias.

Analyses with <2 studies providing data are not shown (maximal walking speed, functional reach (back, left, right), and range of motion are excluded).

Higher value indicates better balance unless otherwise stated.

++ Mean difference (95% CI) was calculated for MMSE and Handgrip test, and standardised mean (95% CI) calculated for all other measures.

Table 2.7
Pre- and post- sensitivity analysis results for observational studies (direct measures of balance)

Comparison or subgroup	No. of studies	N	Effect size	Heterogeneity
*1 Somatosensory Organisation Test (Somatosensory. ratio). ++	3	139	0.90 (-0.58, 2.38)	81%
1.1 Somatosensory Organisation Test (Somatosensory. ratio). ++	2	63	0.16 (003, 0.29)	0%
*2 Somatosensory Organisation Test (Visual ratio). ++	3	139	-2.71 (-3.99, -1.44)	100%
2.1 Somatosensory Organisation Test (Visual ratio). ++	2	63	0.13 (0.03, 0.22)	40%
*3 Somatosensory Organisation Test (Vestibular ratio). ++	3	139	-0.02 (-0.04, 0.00)	0%
3.1 Somatosensory Organisation Test (Vestibular ratio). ++	2	63	-0.02 (-0.04, 0.00)	0%
*4 Static total body stability eyes open (m). Low value indicates better balance.	3	302	-0.37 (-0.74, 0.01)	57%
*5 Static total body stability eyes open (cm^2). Low value indicates better balance.	4	231	-0.89 (-2.11, 0.33)	93%
5.1 Static total body stability eyes open (cm^2). Low value indicates better balance.	2	145	0.34 (-0.25, 0.94)	66%
*6 Static total body stability eyes open (velocity) (cm/s). Low value indicates better balance.	3	161	-1.55 (-3.35, 0.25)	95%
6.1 Static total body stability eyes open (velocity) (cm/s). Low value indicates better balance.	2	135	0.07 (-0.29, 0.43)	2%
*7 Static total body stability eyes closed (velocity) (cm/s). Low value indicates better balance.	3	161	-1.67 (-3.50, 0.16)	95%
7.1 Static total body stability eyes closed (velocity) (cm/s). Low value indicates better balance.	2	135	-3.05 (-9.53, 3.43)	2%
*8 Static ML stability body angle (degrees). Low value indicates better balance.	2	96	-0.12 (-0.52, 0.28)	0%
*9 Static AP stability body angle (degrees). Low value indicates better balance.	2	96	-0.11 (-0.75, 0.53)	60%
*10 Dynamic AP stability (forward) (angle ˚). Low value indicates better balance.	2	72	0.01 (-2.19, 2.22)	94%

Table 2.7 (continued)
Pre- and post- sensitivity analysis results for observational studies (direct measures of balance)

Comparison or subgroup	**No. of studies**	**N**	**Effect size (95% CI)**	**Heterogeneity**
*11 Dynamic Loss of Stability (max excursion) (%). Low value indicates better balance.	2	68	1.09 (0.57,1.60)	0%
*12 Dynamic Loss of stability (directional control) (%). Low value indicates better balance.	2	68	1.02 (0.47, 1.58)	11%

Note:

Data are shown in the table for 12 variables. For some variables there are two sets of data, the first set of data identified with * includes all available data, whereas the second set of data excludes studies at high risk of bias.

Higher value indicates better balance unless otherwise stated.

++ Mean difference (95% CI) was calculated for SOT visual, vestibular and somatosensory ratios, and standardised mean (95% CI) calculated for all other measures.

Table 2.8
Study characteristics of included intervention studies

Study Author	Study Design	Study Population	Physical Activity Measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, Country, setting, consent	**More active (MA) versus less active (LA)**	**Measure, Duration, Intensity**	**Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)**	**Cognitive**	**Sensory**	**Other**	
Paillard et al., 2004	-RCT -Baseline & post 12 weeks -No follow-up -Randomised but not specified -Conflict of interest: N/k -Source of funding: N/k	-N: 21 -Age: 66.15±2yrs (63-72yrs) -0% women -Weight(kg): 74.8±6.7 -Height(m): 1.71±0.05 -Community setting -Written informed consent	Intervention group: 3 months walking programme Control: no walking programme	Baseline measure: n/k MA group: up to 5 hrs of walking per week for 3 months LA group: up to 3 hrs per week no walking programme	**Indirect measure** **(G)** Walking speed (preferred) (velocity - m/min) **Direct measure** *Static balance test* Total body stability (eyes open/closed) (sway distance -- mm; sway area -mm²; speed variation; ratio - EO/EC*100) AP & ML stability (eyes open/closed) (distance - mm; sway area - mm²) *Dynamic balance test* ML stability (eyes open/closed) (position˚; amplitude˚; spectral energy- %	n/a	n/a	n/a	12-week walking programme can improve postural control whilst moving but not when static.

Table 2.8(continued)
Study characteristics of included intervention studies

Study Author	Study Design	Study Population	Physical Activity Measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, Country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Santos Mendes et al., 2011	-RCT -Baseline & post 4mths stratified by sex & randomised -No-follow-up -Conflict of interest: N/k -Source of funding: N/k	-N: 30 -Age 68.7±3.5yrs -60% women -Weight(kg): 66.9 -Height(m): 1.69 -Community setting	Intervention group: 4 month walking prog. Control: no PA	MA group: 1 hour per week for 4 months LA group: no PA	**Direct measure** *Static balance test* Total body stability (8 positions) (Static Balance Index) *Dynamic balance test* Total body stability (2 tests - hurdle obstacle; sit down and stand up from chair) (Dynamic Balance Index)	n/a	n/a	n/a	Walking is beneficial to dynamic and static balance.
Wayne et al., 2014	-RCT -3 time points: Baseline, 3 months, 6 months -No-follow up -Recruitment: N/k -Conflict of interest: N/k -Source of funding: N/k	-N: 60 -Age: 64.19±7.72yrs (50-79yrs) -67% women -White: 92%; Non-Hispanic: 98% -Education: 17±3yrs BMI (kg/m²): 26.5±5.5 -Boston, US -Community setting	MA group; Tai chi expert 6 months tai chi LA group: Tai chi naïve	PASS MA group: 4.0±2.0 (intensity/mins per week) LA group: 4.0±2.0 (intensity/mins per week	**Indirect measure** **(FU)** Timed Up & Go (time to complete -s) **(FU)** Single leg stance (balance time - s) **Direct measure** *Static balance test* Total body stability (eyes open/close) (sway velocity - mm/s; sway area - mm²) Dynamic balance test AP & ML stability (EO/EC) (sway velocity - mm/s)	MMSE (30 items)	n/a	n/a	MA group had no change on COP & some increase in body sway correlated to practice hours.

Table 2.8(continued)
Study characteristics of included intervention studies

Study Author	Study Design	Study Population	Physical Activity Measure (type, level)		Outcome measures of balance				Main Finding
		N, Age (mean, SD & range) % female, race, ethnicity, height (m), weight (kg), BMI, education, Country, setting, consent	More active (MA) versus less active (LA)	Measure, Duration, Intensity	Neuromuscular (Gait (G); Strength (S); Functionality (FU); Flexibility (FL)	Cognitive	Sensory	Other	
Yang et al., 2007	-RCT Baseline, 2-month, 6 months -No follow-up -Randomisation program for 4 locations -Conflict of interest: N/k -Source of funding: N/k	-N: 49 -Age: 80.55±8.49yrs (60-97yrs) -80% women -Retirement home (76%)	MA group: 2 months Tai chi LA group: no tai chi	Measure: n/k MA group: 3 hrs tai chi per week for 2 months LA group: usual activity 3.67±2.38hrs per week	**Indirect measure** **(FU)** Berg Balance (baseline only) **Direct measure** Sensory Organisation Test (somatosensory, visual & vestibular ratios) Base of support (area - cm^2; feet opening angle ˚)	n/a	n/a	n/a	MA group have better SOT vestibular results and greater Base of Support measures but no differences for SOT visual ratios or feet opening angle between groups.

Table 2.9
Summary of the indirect balance outcomes reported in intervention studies

System	**Neuromuscular**				**Cognitive**
	Gait	**Functionality**			**Cognition**
Indirect measures	Max. walk speed (preferred)	TUG	Single leg stance	Berg Balance Scale	MMSE
Paillard et al., 2004	X				
Wayne et al., 2014		X	X		X
Yang et al., 2007				X	

Table 2.10
Summary of the direct balance outcomes reported in intervention studies

Direct measures	**STATIC (Force platform & sway indicators)**	Static balance index	Total body stability (eyes open/closed) (velocity)	Total body stability (distance)	Total body stability (area)	Romberg quotient	AP & ML stability (eyes open/closed) (distance)	AP & ML stability (eyes open/closed) (area)	**DYNAMIC (Force platform & sway indicators)**	Dynamic balance index	Elliptical are (eyes open/closed) (area)	AP & ML stability position/amplitude/spectral energy	**Other**	SOT
Santos Mendes et al., 2011	**X**	X							**X**	X				
Paillard et al., 2004	**X**			X	X	X	X	X	**X**			X		
Wayne et al., 2014	**X**		X		X				**X**		X	X		
Yang et al., 2007									**X**					X

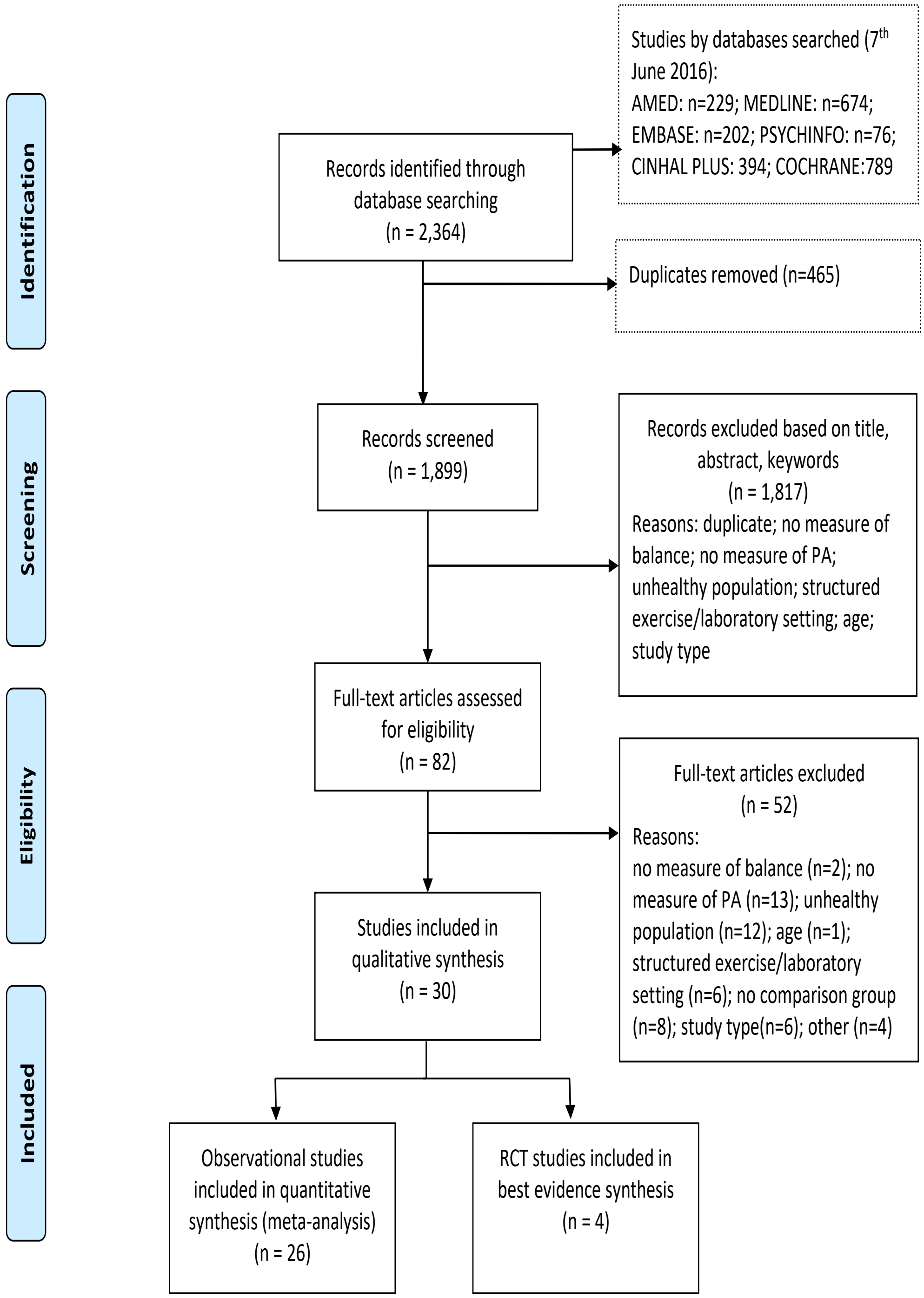
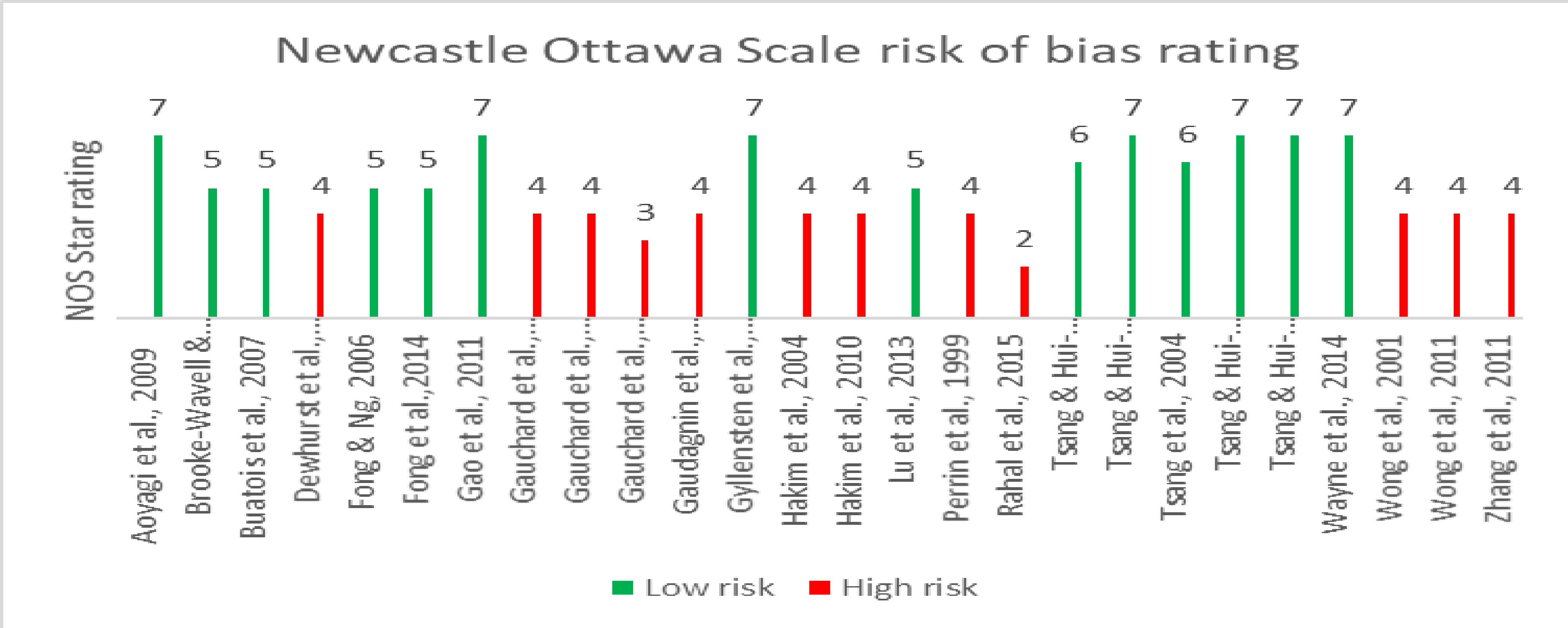


Figure 2.1
PRISMA flow diagram

(source: Moher et al., 2009)



(Risk of bias ratings are: ≤4 stars=high risk; ≥5 stars=moderate to low risk)

Figure 2.2
Risk of bias of observational studies using Newcastle Ottawa Scale

	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants and personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Other bias
Felipe et al., 2011	−	−	−	?	?	+	−
Paillard et al., 2004	−	?	−	−	+	+	−
Wayne et al., 2014	+	?	−	+	+	−	−
Yang et al., 2007	?	+	−	+	−	+	−

Figure 2.3
A summary table of review authors' judgements for each risk of bias item for RCT studies

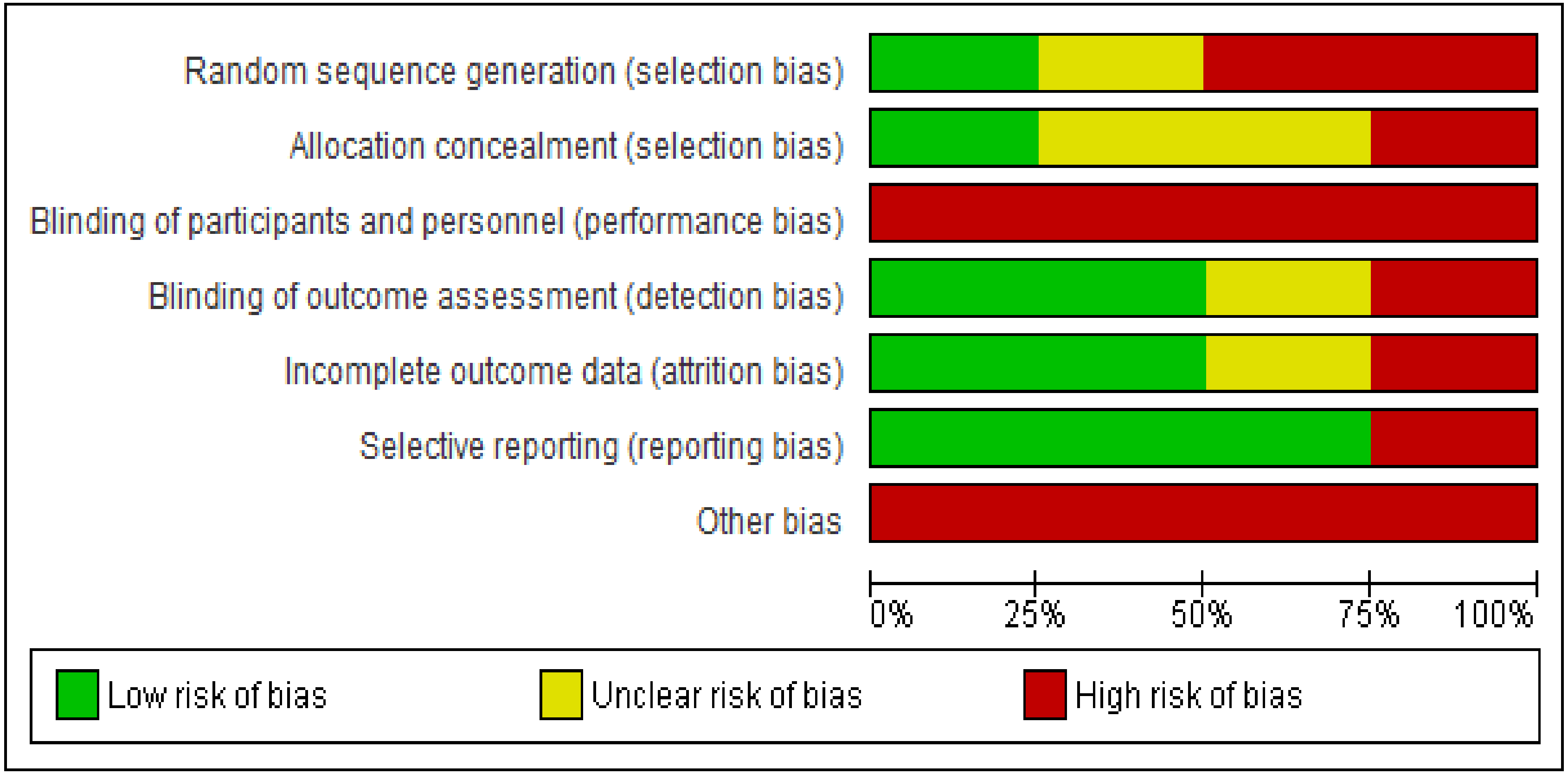


Figure 2.4
A plot of the distribution of review authors' judgements across RCT studies for each risk of bias item

Chapter Three:

The development of the predictive model of free-living physical activity (PA) and balance in an older adult community-dwelling population (≥50 years) using the Irish Longitudinal Ageing study (TILDA)

3.1 Abstract

Background

Poor balance can cause injury, disability, and death in older adults, and so there is a need to improve fall prevention in health care. However, evidence supporting the benefits of physical activity (PA) for balance performance lack methodological robustness, and free-living PA may require a longer period of study than that offered by clinical trials to realise the benefits. Additionally, no gold standard method of measuring balance using a combination of indirect measures exists.

Methods

PA and balance measures from the Irish Longitudinal Ageing study (n=8,504 participants) collected over two years were analysed using a Structural Equation Modelling (SEM) approach to firstly identify an appropriate model of balance, and then to understand and explain the patterns of change in balance and free-living PA over time controlling for other covariates.

Results

The indirect measures of timed up and go test, handgrip test, Mini Mental State Exam, as well as self-rated measures of vision, hearing and steadiness provide a comprehensive model for balance assessment in an older adult community dwelling population in the Republic of Ireland. PA improves balance in the short-term (Est=-0.10, SE=0.12), and cumulatively over two years (Est=-0.13, SE=0.02)

in older adults. Additionally, medication, alcohol consumption, sex, age, fear of falling, education, pain, and problems performing activities of daily living (ADL) were significant risk factors for balance.

Conclusion

This study provides a novel, validated model of balance, and shows that free-living PA benefits for balance are realised both in the short-term and cumulatively over two years in older adults.

3.2 Introduction

There is a lack of methodologically robust evidence to support the benefits of PA on balance performance in older adults (≥50 years). Additionally, a systematic review (Chapter two; McMullan et al., 2018) suggests that there is limited evidence from clinical trials of low moderate quality (four studies), and observational studies of moderate quality that free-living PA benefits balance in community-dwelling healthy older adults (≥50 years). Therefore, more methodologically robust longitudinal studies are needed to explore the association between PA and balance in older adults (≥50 years).

Additionally, Chapter one (section 1.2.2) highlights that despite the focus of PA guidelines on MVPA, older adults are more likely to engage in LPA such as walking for leisure, occupational or transportation purposes. Research suggests that LPA may require a longer time of study than that afforded by clinical studies and may be more appropriate for older adults because there is less risk of injury and can be incorporated into everyday living activity (Bauman et al., 2016; Hallal et al., 2012; Hoffmann et al., 2016; Morris & Hardman, 1997). Consequently, exploring the effects of free-living PA, that is not just MVPA, on balance over time are warranted.

Furthermore, Chapter one (section 1.3.2) highlights the complexity of balance assessment and identified that robust assessment requires neuromuscular, cognitive, and sensory systems to be collectively assessed (Horak, 1995; 2009; Mancini & Horak, 2010; Sibley et al., 2015). However, studies assessing the effects of PA on balance do not measure balance comprehensively across those body systems required for balance with most including neuromuscular measures of balance (19 of 30 studies), a smaller proportion including cognitive measures (ten of 30 studies), and even less including sensory measures (three of 30 studies) (Chapter two, section 2.5.2; McMullan et al., 2018). Additionally, whilst multiple indirect measures (functional tests) were identified as more appropriate for balance assessment due to low cost and ease of implementation (Graham et al., 2008; Horak, 1995; 2006; Sibley et al., 2015; Winter 1995), guidance on which combination of indirect measures provide a robust assessment of balance is lacking (Vieira et al., 2016). Therefore, exploring an appropriate composite measure of balance using multiple indirect or functional measures across different body systems is needed to ensure comprehensive balance measurement (Horak, 1995; 2009; Mancini & Horak, 2010; Sibley et al., 2015).

The inclusion of multiple measures to develop a composite measure of balance involves the use of both objective and self-reported measures that may be subject to measurement error (Chapter one, section 1.3.3). For example, self-reported measures may be influenced by health status, mood, depression, anxiety, or cognitive ability, social desirability, or recall issues (Dyrstad et al., 2014; Murphy, 2009; Saelens et al., 2012). Objective measures may also be subject to bias due to differences between the measurement properties of instruments used, and different protocols for data collection (Bauman et al., 2006; Doherty et al., 2017; Dishman et al., 2001; Haskell & Kiernan, 2000). Therefore, the reliability and the validity of the measures needs to be considered (Borsboom, 2006).

As a result, this study will use data from TILDA, a study of ageing, that provides data relating to PA and balance across two time points from a large sample of community-dwelling participants (≥50 years) (8,504 participants). Also, a SEM approach enables the composite measure of balance to be modelled using multiple indirect measures across neuromuscular, cognitive and sensory systems to confirm whether they can be attributed to the composite measure of balance, and minimises measurement error (Bollen, 1989; McCoach et al., 2007; Raykov & Marcoulides, 2000). Therefore, the methodology and statistical approach used in this study will increase the robustness of the findings and increase our understanding of the effects of free-living PA on a comprehensive measure of balance in community-dwelling older adults (≥50 years).

3.3 Aims and objectives

The overall aim of this study is to develop a predictive model of free-living PA and balance to increase our understanding of the effects of free-living PA for fall prevention in older community-dwelling adults (≥50 years). The following objectives were identified to determine the most appropriate and robust model:

- To identify appropriate indirect measures (functional tests) of balance from the TILDA data that assess all the body systems required for balance (neuromuscular, cognitive, and sensory systems).

- To assess the validity of these indirect measures to robustly provide a composite measure of balance in community-dwelling older adults (≥50 years).
- To understand the effects of free-living PA on balance performance over time in community-dwelling older adults (≥50 years).

3.4 Methods

3.4.1 TILDA Study design

TILDA was chosen for secondary data analysis in this study as it includes a large representative prospective cohort of community-dwelling older adults (≥50 years) in the Republic of Ireland (RoI) (Kearney, Cronin and O'Regan, 2011). It includes relevant balance and PA measures across two points in time over two years for 8,504 participants. Additionally, as the aim of this study is to develop a predictive model of PA and balance, the model must be tested using data from other longitudinal studies (Chapter six). TILDA adheres to the Gateway to Global Ageing Initiative (https://g2aging.org/) which improves the harmonisation of the outcomes, therefore improving the comparability of results across other ageing studies (Appendix VIII).

In brief, the sampling frame used in TILDA was the Irish Geodirectory, a listing of residential addresses from which a clustered sample of addresses was chosen and stratified according to area level socioeconomic status and geographical location. Addresses were selected within each geographic cluster, and all household residents ≥50 years along with their spouses/partners were eligible to participate (Kearney et al., 2011). Data collection included (i) a computer-assisted personal interview (CAPI) that included detailed questions on sociodemographic, wealth, health, lifestyle, social support and participation, use of health and social care and attitudes to ageing; (ii) a self-completed questionnaire; and (iii) a detailed health assessment carried out by qualified and trained research nurses that included cognitive, cardiovascular, mobility, strength, bone and vision tests.

This study uses the available data collected in 2009, and two years later in 2011. In total, the household response rate was 62% for wave one and 86% for wave two. In wave one a total of 8,504 completed the CAPI, of which 8,175 participants were aged ≥50 years, and in wave two a total of 7,455 participants completed the CAPI, of which 7,145 were ≥50 years (6,995 of these participants provided measures across both waves one and two). The health assessment was completed by 5,897 participants in wave one (85.4% in the health assessment centres and 15.6% in their own home) and 7,455 participants in wave two.

The data were provided free of charge through an online application process for the purposes of this analysis by the Irish Social Science Data Archive (ISSDA) at University College Dublin (http://www.ucd.ie/issda/data/tilda/) and the Interuniversity Consortium for Political and Social Research (ICPSR) at the University of Michigan (http://www.icpsr.umich.edu/icpsrweb/ICPSR/studies/34315). Ethical approval for TILDA was obtained from the Trinity College Dublin Research Ethics

Committee, and all participants provided written informed consent. Ethical approval was received from the Filter Committee (Ulster University in January 2017) for the purposes of this study (Appendix IX).

3.4.2 Secondary data analysis approach

The benefits of the Structural Equation Modelling (SEM) approach are highlighted in Chapter one (Table 1.3.3), but in summary using this approach to analyse the TILDA data enables a robust and appropriate predictive model of PA and balance to be developed. Firstly, SEM models can use both unobserved and observed variables, thus enabling the composite measure of balance to be modelled using multiple observed indirect measures of balance across neuromuscular, cognitive and sensory body systems (McCoach et al., 2007). Additionally, within SEM, confirmatory factor analysis (CFA) tests factorial invariance to ensure that the same latent variable of balance is assessed at each time of measurement, thus improving the consistency and reliability of the results for the population being studied (Ferrer et al., 2009; Hoyle, 1995). Also, by examining the structural relationships between the measured variable of PA and the composite measure of balance (Hoyle, 1995; Stoel et al., 2003), estimates of individual trajectories of change can be assessed (Jung & Wickrama, 2008), thus providing an understanding of how balance and free-living PA in older adults (≥50 years) change over two points in time (Jung & Wickrama, 2008). Furthermore, due to the length of study (two years) there may be missing data due to drop out or mortality, and within SEM, a robust form of Maximum Likelihood Estimation (MLE) uses a model-based estimation strategy for missing data thereore reducing standard errors.

3.4.3 Measures

The appropriate variables for balance, free-living PA, as well as covariates or exogenous variables (i.e. age, sex, education, falls, fear of falling, medication, pain, alcohol, sleep, and ADL) (Figure 3.1) that were provided at both wave one and two, were identified in the TILDA data using a combination of the Derived Variables Codebook for wave one and two (2016) and the TILDA Release Guide (2016). These were then prepared for analysis in M-plus (version 7.4) in accordance with the M-plus user manual (Muthén & Muthén, 1998-2017) as detailed in Table 3.1 (page xlvii) and outlined below.

3.4.3.1 Balance

Balance requires the contribution from cognitive, neuromuscular, and sensory systems and therefore six indirect measures (functional tests) (available at both waves) across multiple body systems were selected from TILDA to develop a composite measure of balance. Although these measures test different body systems, they also test a different aspect of balance, and so collectively provide a comprehensive assessment of balance. A low score on balance indicates good balance using the scoring system in this study. Descriptive statistics for each measure of the six indirect measures used to devleop the composite measure of balance are outlined in Table 3.2 (page lviii).

Sensory System

Vision and hearing decline with age and are associated with falls and ultimately a decline in quality of life (Chia et al., 2006). Research suggests an association between conversational hearing and vestibular disfunction in older adults (Lin & Ferrucci, 2011); poor vision and increased sway in older adults (Black et al., 2008); and proprioceptive feedback issues and poor balance in older adults (Pickard et al., 2005). Self-rated measures of vision and hearing are reported to be good indicators of actual vision and hearing when compared with objective measures (El-Gasim et al., 2013; Valete-Rosalino & Rozenfeld, 2005). The following questions were used to assess vision and hearing: “Is your eyesight (using glasses or corrective lenses)?”; “Is your hearing (with or without a hearing aid)?”, excellent (1), very good (2), good (3) fair (4) or poor (5)?” A low score indicates good vision or hearing.

Cognitive System

Cognitive ability was assessed using a summary score of the Mini Mental State Exam (MMSE) which measures attention, concentration, memory, language, visio-constructional skills, calculations, and orientation (maximum score of 30) (Folstein et al., 1975). Age is a risk factor for cognitive degeneration which can affect balance (Forbes et al., 2015). Research suggests that cognitive decline is associated with poor balance in older adults (Tangen et al., 2014). A high score indicates good cognitive function.

Neuromuscular System

Strength was assessed using the highest score from three tests on the dominant hand (Kg) in the handgrip test, a predictor of overall body strength which is required for good balance (Bohannon & Shaubert, 2005). A hydraulic hand dynamometer was used (Fabrication, White Plains, NY, USA). A high score indicates good strength.

Strength, mobility and gait speed were assessed using the objective measure of the Timed Up and Go test (TUG) (seconds) (Podsialdlo & Richardson, 1991). Strength and mobility are required for balance and gait patterns change with age where greater gait variability is predictive of falls (Maki, 1997). The time to rise from a chair, walk three meters at normal pace, turn, walk back, and sit down was recorded. Walking aids were allowed, and no instruction was provided about the use of participant's arms. The time taken from the command "Go", to when the participant sat with their back resting against the back of the chair was recorded using a stop watch. A low time indicates a good score.

Steadiness was assessed using three questions: "when standing..." "when getting up from a chair..." and "when walking...do you feel?", (1) very steady; (2) slightly steady; (3) slightly unsteady; (4) very unsteady. The summed score of steadiness is shown to be a predictive and reliable measure of disability and falls in community-dwelling older adults (Clark, Callahan & Counsell, 2005; Lindenberger et al., 2003). Results from a factor analysis confirmed that all three questions were closely correlated (Est=0.94, 0.95, and 0.88 respectively, therefore the summed score of all three questions was used for analysis. A low score indicates good steadiness.

3.4.3.2 Free-living physical activity (PA)

The International Physical Activity Questionnaire (IPAQ) (short-form), a validated measure of PA was used to assess PA levels (Craig et al., 2003; Hallal et al., 2012). IPAQ is a self-reported measure of time spent on different activity levels (vigorous/moderate/walking) over the last seven days. The time spent on activity level is weighted based on energy requirement giving a total score of MET-minutes per week (the metabolic energy spent on activities multiplied by the amount of time spent doing them). To improve the validity of the measure, the score was corrected for measurement error (Borsboom, 2006; Saris & Stronkhorst, 1984) using the inter correlation coefficient (ICC) of 0.76 (Craig et al., 2003). A high score indicates higher levels of PA. Descriptive statistics for the measure of free-living PA are outlined in Table 3.2 (page lviii).

3.4.3.3 Covariates/exogenous variables

Factors affecting balance that are not caused by other variables in the model were included, and where measures included multiple indicators (alcohol, sleep and pain) a factor model was used to assess internal validity and based on the results appropriate questions were included.

Demographics

The variable of age was used (mean=74.3yrs), and sex was recorded as female (42%) or male (58%). Highest education level was recorded where primary education is <11 years of full-time (27%); secondary education is 11-13 years of full-time (60%); and tertiary education includes diploma/degree/higher (13%). Primary and secondary education variables have been used for analysis.

Lifestyle and health

Fear of falling is measured using one question “Are you afraid of falling?” with two response options of “yes” or “no”. Thirty percent of participants had a fear of falling. Seventy-four percent of participants had fallen in the last year. Eighty percent of participants were taking medication.

A factor model indicated that one of the two questions relating to pain had a high factor loading (Est=1) and so one question regarding how pain affected usual day to day activities was included (55% of participants confirmed that pain affected their everyday activities).

A factor model showed that all three questions used to assess alcohol consumption had high and equivalent factor loadings (Est=0.93, 0.90, and 0.95 respectively) and so an equally weighted summed index was used: (1) “Have people annoyed you by criticising your drinking?"; (2) "Have you ever felt guilty about drinking?"; and (3) "Have you ever felt you needed a drink first thing in the morning to steady your nerves or to get rid of a hangover?" (91% experienced alcohol issues).

A factor model identified that two of three questions assessed sleep quality (Est=0.070 and 0.82 respectively): (1) "How often do you have trouble falling asleep?" (18%), and (2) "How often do you have trouble with waking up too early and not being able to sleep? (40% of participants experienced poor sleep quality). Participants were asked if they had difficulty performing any of six activities of daily life (ADL) such as dressing, walking across a room, bathing or showering, eating, getting in or out of bed, and using the toilet, and 85% experienced problems.

3.5 Statistical analysis

Using a SEM approach, analysis was carried out in Mplus (version 7.4; Muthen & Muthen, Los Angeles, CA). A complex analysis option in Mplus was used to account for clustering and stratification. A Maximum Likelihood Estimation method with robust standard errors (MLR) was used.

Model fit was evaluated using a Root Mean Square Error of Approximation (RMSEA) ≤0.05 with an upper limit (90% CI) ≤0.08; a Comparative Fit Index (CFI) ≥0.95; a Tucker Lewis Index (TLI) ≥0.95; and a Standardised Root Mean Square Residual (SRMR) ≤0.08 (Hoyle, 1995). Where the levels of fit indices were not achieved, the modification indices were examined, and where appropriate, adjustments made (Appendix X).

Confirmatory factor analysis (CFA) explored whether handgrip strength, Vision, Hearing, Steadiness, TUG, MMSE (Figure 3.1) could be attributed to the composite measure of balance at both waves (configural invariance); whether each measure demonstrated equal relationships with balance across time (metric invariance); and whether differences over time in balance are due to true change in the underlying measures (scalar invariance). The mean difference of the composite measure of balance was then examined by fixing the mean of balance at zero and exploring the change at wave two (Sorbom, 1974).

After determining whether balance was invariant over time, the relationships between balance and free-living PA were examined by allowing free-living PA to have a direct effect on baseline balance (Figure 3.1).

Finally, each covariate/exogenous variable was introduced into the model and regressed onto free-living PA at baseline. A direct effect was introduced from the exogenous measures to composite measure of balance if model fit indices indicated that the model did not adequately describe the data. A series of sensitivity analyses were conducted to check the effects of the reducing numbers of responses in some of the exogenous variables, but no substantive changes were found in the parameter estimates. The final model developed for the model of free-living PA and balance using the TILDA data is included in Appendix XI.

3.6 Results

Baseline and two-year descriptive statistics for the observed variables of balance and physical activity are shown in Table 3.2 (page lviii). A summary of the results is shown in Table 3.3 (page lix) and described below.

3.6.1 Balance

Standardised factor loadings indicated that all six indirect measures (MMSE (Est=-0.33, S.E.=0.04), vision (Est=0.27, S.E.=0.02), hearing (Est=0.23, S.E.=0.03), handgrip strength (Est=-0.22, S.E.=0.03), TUG (Est=0.71, S.E.=0.04) and steadiness (Est=0.86, S.E.=0.04), had a statistically significant relationship with the composite measure of balance. A residual correlation between vision and hearing was introduced because these measures showed a variance not explained by balance (Figure 3.2, page lxiii). A series of successive restrictions on the factor loadings for each measure of balance (metric invariance) can be assumed for each factor loading at both waves, showing that each measure demonstrated equal relationships with balance across time. Scalar invariance could not be assumed for the measures of balance excluding MMSE, demonstrating partial invariance (Table 3.4, page lxi shows a summary of the fit statistics for configural, metric,and scalar invariance analysis). Balance at baseline and wave two was highly correlated, (Est=0.98). The mean difference score between baseline and wave two balance shows that balance declined after two years (Est=-0.67). In other words, a one-unit change in the baseline score results on average in a change of only 0.676 units and not a value of one, which it would be if no change had occurred. This amounts to a reduction of approximately 0.2514 (normal distribution table) of one unit across this two-year period *(using normal distribution table to work out the % differences).*

3.6.2 Free-living physical activity (PA)

Baseline free-living PA had a statistically significant direct effect on free-living PA at wave two (Est=0.40). Based on modification indices a direct effect was introduced for handgrip strength on baseline free-living PA (Est=-0.4) and wave two PA (Est=0.1).

3.6.3 Free-Living physical activity (PA) and balance

3.6.3.1 Direct effects

Our model assumes that free-Living PA influences balance, and analysis found that baseline free-living PA has a statistically significant effect on baseline balance, where an extra MET-minute of free-living PA per week improves balance by -0.10 SDs or 4% (normal distribution table) (Figure 3.1, path B, page lxii), but had no statistically significant effect ($p>0.05$) on wave two balance (Est=0.04) (Figure 3.1, path C, page lxii). The data for free-living PA and balance are at the same time point (baseline) and so it was not possible to also test the effect of baseline balance on free-living PA, because there are no independent uncorrelated predictors for balance or free-living PA. Baseline balance was shown to have a statistically significant positive effect on wave two free-living PA (Est=-0.14) (Figure 3.1, page lxii, path E).

Wave two free-living PA has a statistically significant effect on wave two balance, where an extra MET-minute of free-living PA per week improves balance by -0.05 SDs or 2% (normal distribution table) (Figure 3.1, page lxii, path D). Figure 3.3 (page lxiv) shows a more detailed representation of the results.

3.6.3.2 Indirect effects

Baseline free-living PA has a statistically significant total indirect effect on wave two balance via the effect of wave two free-living PA (Figure 3.1, page lxii, path F, D); via the effect on baseline balance (Figure 3.1, page lxii, path B, A); and via the effect of baseline balance on wave two free-living PA (Figure 3.1, page lxii, path B, E, D) (Est=-0.13), where an extra MET-minute per week of free-living PA improves balance by -0.13 SDs or 5% over two years.

3.6.3.3 Covariates

Gender (Est=-1.28), medication (Est=-0.98), and Activities of Daily Living (ADL) (Est=-0.2) had a statistically significant effect on free-living PA, and because free-living PA indirectly affects balance, then an indirect effect on balance. For example, females, those taking medication, or with any ADL impairments engaged in less free-living PA, resulting in poorer balance. Additionally, increased age (Est=0.15), fear of falling (Est=1.13), lower education (primary: Est=1.09;

secondary: Est=0.7), pain (Est=-0.23), higher alcohol consumption (Est=-0.31), and problems performing ADL (Est=2.12), over and above their effect on free-living PA (i.e. an independent effect), were found to adversely directly affect balance. Sleep quality and a history of falls were not significant for free-living PA or balance.

3.7 Discussion

3.7.1 Principal findings

CFA analysis supports that multiple indirect functional measures such as MMSE (cognitive); TUG, handgrip test, and steadiness (neuromuscular); and vision and hearing (sensory) collectively provide a composite measure of balance assessment in a community-dwelling older population (≥50 years) in the RoI (Horak, 1995; Sibley et al., 2015).

Additionally, whilst only partial invariance was found and as a result, the mean difference between balance at baseline and two years should be viewed with caution, the mean suggests that balance declines with age (Horak, 2006). The findings also suggest that, free-living PA, the activity of everyday living, prevents balance decline. Therefore, PA that is not just exercise, can improve or maintain balance in older age, thus suggesting that free-living PA has a cumulative effect on balance over a two-year period.

The findings also support existing research that suggests that being female (WHO, 2007); using medication (Blake, 1988); and having a problem performing an ADL (Bandeen-Roche et al., 2015; Tak et al., 2013) results in lower activity levels which adversely affects balance. Additionally, increased age, fear of falling (Bandeen-Roche et al., 2015), lower education (Chen et al., 2015; Preston, et al., 1998); pain (Marmot et al., 2002); and high alcohol consumption (WHO, 2007) are also confirmed as important risk factors for poor balance. The findings do not support that sleep quality or history of falls are significant for either PA or balance.

3.7.2 Methodological quality

Chapter two highlighted that research exploring the effects of free-living PA and balance were in the main, of low to moderate quality, and this study addresses some of the methodological issues highlighted. For example, this study uses a large representative sample of data over a two-year period from a robustly designed longitudinal study of ageing (TILDA), thus the ecological validity is high, and the robustness of the conclusions drawn strengthened. Also, the SEM statistical approach adopted minimises measurement error and validates the measures for older adults living in RoI, therefore further strengthening the findings. Additionally, the analysis uses MLR, thus allowing all participant data to be used, and therefore reduces the bias of the results (Enders, 2013).

3.7.3 Strengths and weaknesses

A key strength of this study is the approach used for balance assessment as it uses multiple indirect measures across neuromuscular, sensory and cognitive body systems to assess balance holistically (Sibley et al., 2015). These indirect measures are also more appropriate for use in a clinical or community setting due to low cost and ease of implantation (Winter, 1995). Additionally, guidance on what combination of indirect measures provide a robust assessment of balance is lacking (Vieira et al., 2016), and this study provides an indication that measures such as MMSE (cognitive), TUG, handgrip strength, and steadiness (neuromuscular) are strong measures for balance, whilst self-reported measures of vision and hearing have a weaker relationship as they may be affected by a method effect due to recall bias (Hassan, 2006), interpretation bias (Mazor et al., 2002), or a bias caused by over- or under-estimation (Ramkissoona & Cole, 2011; El-Gasim et al., 2013; Sakurai et al., 2013).

Additionally, as previously highlighted in chapters one and two, the effects of PA that is not just exercise, may require a longer duration of study than that afforded by clinical trials (Morris and Hardman, 1997). This study uses longitudinal data over a 2-year period to understand the effects of PA, that is not just exercise, on balance in older adults, and so more fully assesses the effects of PA on balance over time.

Also, whilst the use of secondary data can provide a more accurate estimate of effect than clinical trials which may include smaller numbers, unrepresentative samples of populations, and shorter timeframes for outcome assessments (Smith et al., 2011), thus providing an opportunity to make recommendations that may be more relevant to policy makers (Cooke & Iwashyna, 2013), the measures included in this analysis as well as the population sample are restricted to TILDA. Also, this study uses only two waves of data that were available at the time of study, and whilst analysis can provide an indication of the amount of change over time, it cannot provide an understanding of the trajectory or the rate of individual change over time (Duncan & Duncan, 2009).

3.7.4 Future research and clinical implications

This study presents a model of balance that can guide balance assessment within clinical practice because it uses multiple indirect balance measures therefore ensuring that balance is comprehensively measured (Horak, 1995; Sibley et al., 2015), and also tests that are already in use within clinical settings (Graham et al., 2008). As a result, our ability to assess risk of falls more accurately, and target interventions more appropriately is enhanced. Future research should consider using an objective measure of vision such as LogMAR (Minimal Angle of Resolution) (Baily & Lovie, 1975) charts and hearing such as the pure tone audiometry test (PTA) to address the methods effect highlighted above (section 3.7.3). Consideration should also be given to exploring the convergent validity between indirect and direct measures to further assess the appropriateness of the model of balance using indirect measures. Additionally, consideration should be given to a trial using the measures identified within clinical practice to ensure their appropriateness in relation to ease and length of time to complete.

In addition, the results show that free-living PA, that is not just exercise, can benefit balance in the immediate term as well as have a cumulative effect over time. Therefore, programmes of activity for older adults may be developed that may not only be beneficial to balance, but also more appropriate to this population given that this population are failing to meet exercise guidelines (Hallal et al., 2012). For example, barriers to exercise guidelines such as poor health, fear for personal security, and lack of interest or time (Schutzer & Graves, 2004) may be overcome if advice includes activities that are carried out as part of everyday living, such as for example walking, or household chores. Additionally, there are both immediate and cumulative benefits of PA on balance in older adults, thus increased activity should be promoted in older adults to ensure the maintenance or improvement in balance in later life.

Furthermore, to generalise the findings, future research should consider using other ageing studies such as NICOLA (Craig et al., 2012). Also, ELSA provides six waves of data from over 12, 099 participants across 12 years and would enable the robustness of the TILDA analysis to be tested and help understand the trajectory or the rate of individual change over time therefore allowing the effects of the intensity of PA on balance to be explored over time (Duncan & Duncan, 2009).

3.8 Conclusion

There is methodologically robust evidence from the analysis of data from the TILDA study that free-living PA improves or maintains balance measures in older adults both in the short-term and cumulatively over a two-year period. It shows that being generally more active in later life can prevent falls in older adults in RoI. The TILDA data included only two time points and so ELSA data which includes data from six timepoints across ten years will further explore the trajectory of change in PA and balance in older adults. Also, an exploration of the best method to develop the composite measure of PA using categorical data from the ELSA study will be developed to improve the robustness of the results. Additionally, to generalise the findings, the robustness of the findings from the TILDA and ELSA analyses will tested using data from the NICOLA study.

Table 3.1

The measures used to explore the association between free-living physical activity (PA) and balance from TILDA across waves one and two

Free-living Physical Activity (PA)

Original variable name	New variable name Wave 1	New variable name Wave 2	Measure unit and details	Balance performance indicator	Reliability reference
IPAQMETmins	IPAQmmw1	IPAQmmw2	Objective measure Summed score in METmins *	High is good	0.76 (Craig et al., 2003)

*TILDA provides physical activity measures in METS minutes which are a measure of energy expenditure. METS are multiples of the resting metabolic rate and a MET-minute is computed by multiplying the MET score of an activity by the minutes performed (www.ipaq.ki.se, 2005). Recommended MET minutes per week is 500 to 1000 METmins (Office of Disease Prevention, 2018).

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from TILDA across waves one and two

Balance measures

Measure	**Original variable name (source) (number of items)**	**New variable name Wave 1**	**New variable name Wave 2**	**Measure unit and details**	**Balance performance indicator**	**Reliability reference**
Mini mental state exam	Wave 1: cogMMSE_ha Wave 2: mmsescr_capi (wave 1 HA; wave 2 CAPI) (1 measure total scores out of 30 at both waves)	MMSEw1	MMSEW2	Objective measure Total score (max. 30)	High is good	Folstein et al., 1975
Hand Grip test	Wave 1: FRgripstrengthND Wave 2: gs005 (HA) (Average of 3 measures of dominant hand both waves Kg)	GripDw1	GripDw2	Objective measure 1 measure for dominant hand (Kg)	High is good	Bohannon & Shaubert, 2005

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from TILDA across waves one and two

Original variable name	New variable name Wave 1	New variable name Wave 2	Measure unit and details	Balance performance indicator	Reliability reference	Original variable name
Time up and Go	Wave 1: FRtugTimeSec Wave2: tug009s (time taken to complete TUG in seconds) (HA) (1 measurement in seconds both waves)	TUGSw1	TUGSw2	Objective measure Time in seconds (1 tEst) (Secs)	Low is good	Shumway-Cook et al., 2001
Self-rated vision	Wave 1 and 2: ph102 Is your eyesight (using glasses or contact lens if you use them) (scale 1-5) (1 measure available both waves)	Vis01w1	Vis01w2	Self-reported measure Scale (1-5)	Low is good	Kaplan & Camacho, 1983; Idler & Angel 1990

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from TILDA across waves one and two

Original variable name	New variable name Wave 1	New variable name Wave 2	Measure unit and details	Balance performance indicator	Reliability reference	Original variable name
Self-rated hearing	Wave 1 and 2: ph108 Is your hearing (with or without a hearing aid) (scale 1-5) (1 measure available both waves)	Hear02w1	Hear02w2	Self-reported measure Scale (1-5)	Low is good	Kaplan & Camacho, 1983; Idler & Angel 1990
Self-rated steadiness	Wave 1 and 2: ph411 When walking, do you feel ph412 When standing, do you feel ph413 When getting up from a chair, do you feel (scale 1-4) (3 measures available across both waves)	Stead1su (summed value)	Stead2su (summed value)	Self-reported Summed score of the 3 questions relating to steadiness was used	Low is good	Clark et al., 2005

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from TILDA across waves one and two

Covariates

Measure	**Original variable name (source) (number of items)**	**New variable name**	**Measure unit and details**	**Note**
Age	Age (years) (CAPI) (1 item)	Re_Age	Years	Original age variable was centred (actual age minus mean age) wave 1 but differences in model results did not vary between using centred and uncentred variable so reverted back to using the original variable
Sex	Sex (male/female) (CAPI) (1 item)	Re_sex	Sex was recoded (reversed female/male) Female=1 Male=0	Recode to align to other variables

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from the TILDA across waves one and two

Measure	Original variable name (source) (number of items)	New variable name	Measure unit and details	Note
Education	Edu_level (CAPI) (1 item)	Edu_prime Edu_second Edu_third 3 new variables were created	1 item each for each variable answered Yes=1 No=0	Creation of 3 variables to create a reference group (i.e. the group used for comparison) Education was used to represent SES instead of the SES variable provided by TILDA as interpretation of categories was not clear (1= employers and managers; 2=higher professional; 3=lower professional; 4=non-manual; 5=manual skilled; 6=semi-skilled; 7=unskilled; 8=own account workers; 9=unknown; 10=farmers; 11=agricultural workers

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from the TILDA across waves one and two

Measure	Original variable name (source) (number of items)	New variable name	Measure unit and details	Note
Fall history	Ph401 Have you fallen in the last year? (yes/no) (CAPI) (1 item)	Re_falls	Original variable recoded (reversed yes/no) Yes=1 No=0	Recoded to align with other variables
Fear of falling	Ph408 Are you afraid of falling? (yes/no) (CAPI) (1 item)	Re_fefall Are you afraid of falling?	Original variable recoded (reversed yes/no) Yes=1 No=0	Recoded to align with other variables

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from TILDA across waves one and two

Measure	Original variable name (source) (number of items)	New variable name	Measure unit and details	Note
Sleep behaviour	Bh201 How likely are you to doze off or fall asleep during the day? Bh202 How often do you have trouble falling asleep? Bh203 How often do you have trouble with waking up too early and not being able to get back to sleep? 1=most of time 2=sometimes 3=never (CAPI) (3 measures)	Sleep2w1 How often do you have trouble falling asleep? Sleep3w1 How often do you have trouble with waking up too early and not being able to get back to sleep?	1=most of time 2=sometimes 3=never High is good	A factor model was investigated and showed that all three questions were not closely correlated. So, it was decided to exclude bh201 as bh202 and bh203 were more closely correlated.

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from TILDA across waves one and two

Measure	Original variable name (source) (number of items)	New variable name	Measure unit and details	Note
Activities of Daily Living	NADLw1 Number of ADL impairments? (CAPI) (1 measure - total number of ADL impairments max. 5)	Re-ADLw1 Number of ADL impairments?	Yes=1 No=0	Original variable recoded from number of impairments to yes or no.
Pain	Ph504 Does pain make it difficult for you to do your usual activities successfully? (yes/no) Ph505 Are you taking any medication for pain? (yes/no) (CAPI) (2 measures)	Pain1w1 Does pain make it difficult for you to do your usual activities successfully?	Yes=1 No=0	A factor model was runusing both questions relating to pain found that the 2 questions were not well correlated, so a summed score was not used, and it was decided to use only pain1w1 because the question relates to pain rather than medication for pain. Medication is covered in the variable re_meds.

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from TILDA across waves one and two

Measure	**Original variable name (source) (number of items)**	**New variable name**	**Measure unit and details**	**Note**
Medication	MDmeds4 Number of regular medications taken? (CAPI) (1 measure)	Re_meds Regular medications taken?	Yes=1 No=0	TILDA also provides the variable PolyMDw1 Five or more medications (MDpolypharmacy) wave 1? Yes/no. It was decided to use MDmeds4 Number of regular medications taken? as we wanted to understand whether any medication would affect balance.
Alcohol consumption	SCQcage1 cut down on drinking? (yes/no) SCQcage2 criticised your drinking (yes/no) SCQcage3 felt bad or guilty about drinking? (yes/no) (SCQ) (3 measures)	Alch1_sum	Summed score of 3 items	A factor model run for 3 items in alcohol and showed the internal consistency of for alcohol was strong (0.929). As a result, as all 3 items are highly inter-related a summed score was used.

Table 3.1 (continued)
The measures used to explore the association between free-living physical activity (PA) and balance from TILDA across waves one and two

Additional variables

Purpose	Variable name	Description
Clustering	**cluster**	The TILDA sample was recruited from households selected in geographic clusters, and when a household was selected every eligible member of that household was invited to participate. Failing to consider the correlation between participants introduced by this sampling design will lead to biased estimates.
Stratification	**stratum**	The selection of geographic clusters was stratified, so that equal numbers of clusters were selected from each of three socio-economic groups. The socio-economic status of a cluster was defined by the proportion of individuals in that cluster. The variable '**stratum**' indicates to which of the three strata the cluster from which each participant was recruited belonged.
Weight	**capiweig**	CAPI weight based on age/sex/education crosstabulation from 2010 QNHS. Applying these weights to analyses yields Estimates that are applicable to the Irish population in 2010. CAPI' weight, to be applied when the whole TILDA sample is included in an analysis.

Table 3.2

Descriptive statistics for the variable of physical activity and the observed variables of balance (wave one and two)

Observed variables	Wave 1	Wave 2
	(N=population; mean (standard deviation); range)	
Balance		
Vision (Likert scale 1-5) (high score is poor)	N=1709; 2.47 (0.99)	N=1529; 2.56 (0.89)
Hearing (Likert scale 1-5) (high score is poor)	N=1709; 2.46 (1.20)	N=1530; 2.60 (1.04)
MMSE (max. score 30) (high score is good)	N=1406; 28.30 (3.86); (range 15-30)	N=1530; 28.54 (3.96); (range 15-30)
Hand Grip test (kg) (high score is poor)	N=1381; 26.05 (106.53); (range 2-65)	N=1412; 29.26 (158.21); (range -98-75)
TUG (secs) (high score is poor)	N=1392; 9.34 (13.25); (range 4.82-63.53)	N=1483; 9.81 (14.43); (range 2-51)
Steadiness (Likert scale 1-5) (high score is poor)	N=1707; 4.43 (4.74)	N=1707; 4.52 (5.04)
Free-living PA measure		
Free-living PA (total METS mins per week) (high score is good)	N=1707; 2.72 (10.19); (range 0-19.28)	N=1709; 2.19 (9.40); (range 0-17.89)

Table 3.3

The effects of free-living physical activity and covariates (sex, age, medication, falls, education, pain, alcohol, fear of falling and activities of daily living (ADL)) on balance

Description	Estimate (Est)[3]	Standard Error (S.E.)	Est./S.E.[1]
Balance & Free-Living Physical Activity			
Balance (wave 1) on Balance (wave 2)	1.07(0.95)	0.05	20.49
Free-living PA (wave 1) on Balance (wave 1)	-0.10(-0.12)	0.02	-4.19
Free-living PA (wave 1) on Balance (wave 2)	0.04(0.04)	0.03	1.39
Free-living PA (wave 2) on Balance (wave 2)	-0.05(-0.05)	0.02	-2.71
Balance (wave 1) on free-living PA (wave 2)	-0.14(-0.13)	0.03	-4.72
Free-living PA (wave 1) on Free-living PA (wave 2)	0.40(0.40)	0.05	7.78
Direct effects of covariates on free-living physical activity			
Sex	-1.28	0.19	-6.55
Age	-0.03	0.01	-3.46
Medication [4]	-0.98	0.28	-3.53
Falls	0.67	0.20	3.42
Education-primary [2]	0.42	0.27	1.56
Education-secondary [2]	0.56	0.24	2.30
Pain	0.15	0.04	3.36
Alcohol	0.06	0.14	0.42
Sleep (2w1)	0.31	0.13	2.41
Sleep (3w1)	-0.12	0.13	-0.92
Fear of Falling	-0.52	0.18	-2.83
ADL	-0.99	0.25	-4.00

Table 3.3 (continued)

The effects of free-living physical activity and covariates (sex, age, medication, falls, education, pain, alcohol, fear of falling and activities of daily living (ADL)) on balance

Description	Estimate (Est)[3]	Standard Error (S.E.)	Est./S.E.[1]
	-	-	-
Age	0.15	0.01	12.25
Medication	0.13	0.10	1.21
Falls [5]	-	-	-
Education-primary [2]	1.10	0.18	6.00
Education-secondary [2]	0.70	0.14	5.02
Pain	-0.23	0.03	-8.32
Alcohol	-0.31	0.08	-3.91
Sleep (2w1) [5]	-	-	-
Sleep (3w1) [5]	-	-	-
Fear of Falling	1.13	0.17	6.53
ADL	2.12	0.30	7.02

Note:

[1] test statistic showing statistical significance where >1.96 at p=0.05 level.

[2] Reference group for education is Education-third level (e.g. university level).

[3] Unstandardised results are reported with standardised estimates in brackets.

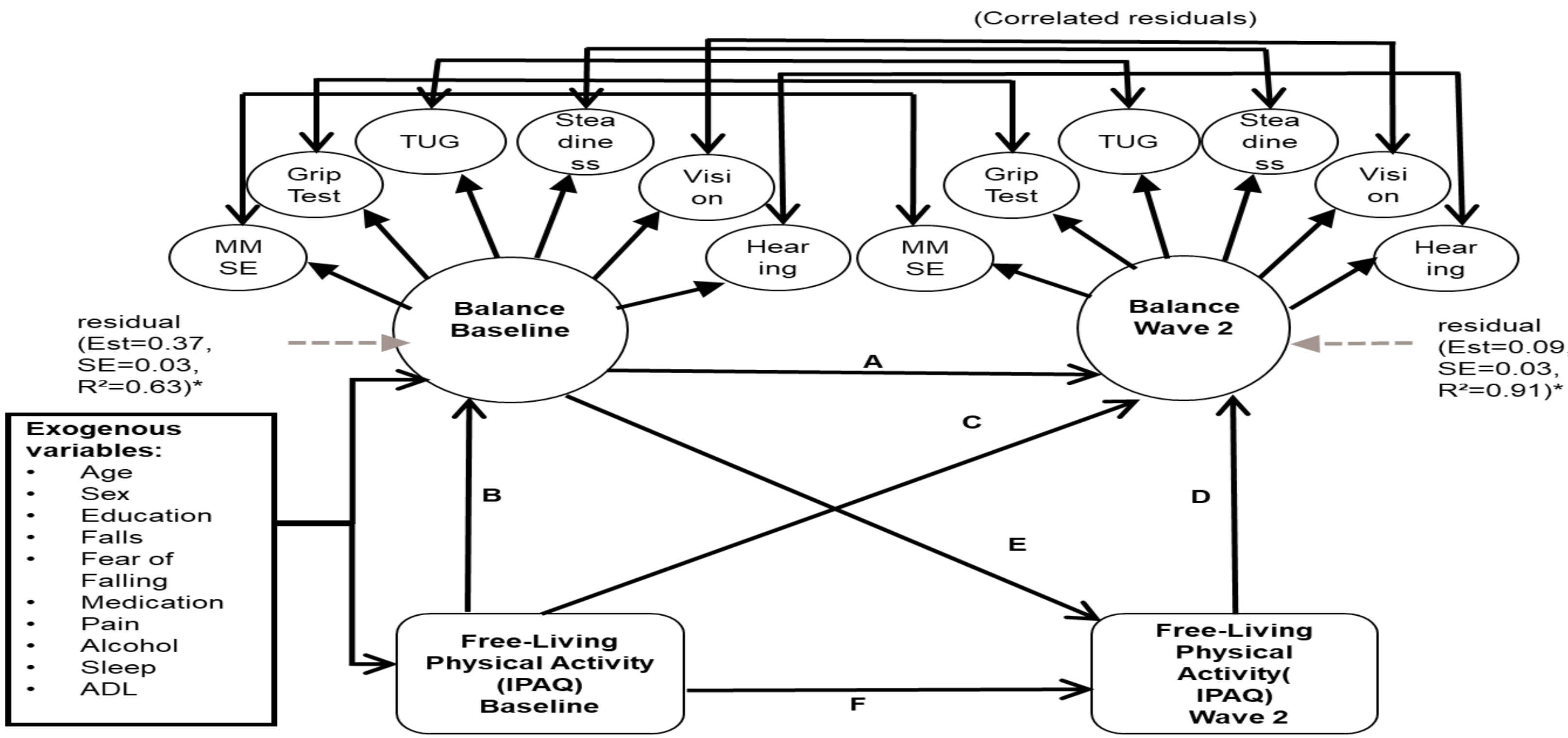
[4] Direct effect of medication on balance is insignificant when controlling for the direct effect on free-living PA.

[5] - indicates that modification indices suggested no direct effect was required.

Table 3.4
Fit statistics for the model of balance at wave one and two, and the model of free-living PA, balance and covariates

Models	Information Criteria		Chi squared			RMSEA[1]		CFI[2]/TLI[3]		SRMR[4]
	Akaike (AIC)	Bayesian (BIC)	value	df	P-value	Estimate	90 % C.I.	CFI	TLI	Value
1.1 Model of balance at wave 1 & 2 (configural variance)	338974.21	339303.48	253.90	43	0.0000	0.03	(0.02, 0.03)	0.97	0.95	0.03
1.2 Model of balance at wave 1 & 2 (metric invariance)	339283.97	339571.20	342.29	49	0.0000	0.03	(0.02, 0.03)	0.96	0.94	0.05
1.3 Model of balance at wave 1 and 2 (scalar invariance)	339117.30	339397.53	313.07	50	0.0000	0.03	(0.02, 0.03)	0.96	0.95	0.04
1.4 Model of Free-living PA, balance & covariates at wave 1 & 2	90876.18	91322.56	503.74	205	0.0000	0.03	(0.03, 0.03)	0.95	0.94	0.04

Note: (Hoyle, 1995).
[1] RMSEA is the Root Mean Square Error of Approximation (≤0.05 with an upper limit (90% Confidence Interval (CI)) ≤0.08)
[2] CFI is the Comparative Fit Index (≥0.95).
[3] TLI is the Tucker Lewis Index (≥0.95)
[4] SRMR is the Standardised Root Mean Square Residual (≤0.08)



(adapted from approach by Best et al. 2015)

Figure 3.1
The relationship between free-living PA and balance over a 2-year period controlling for covariates

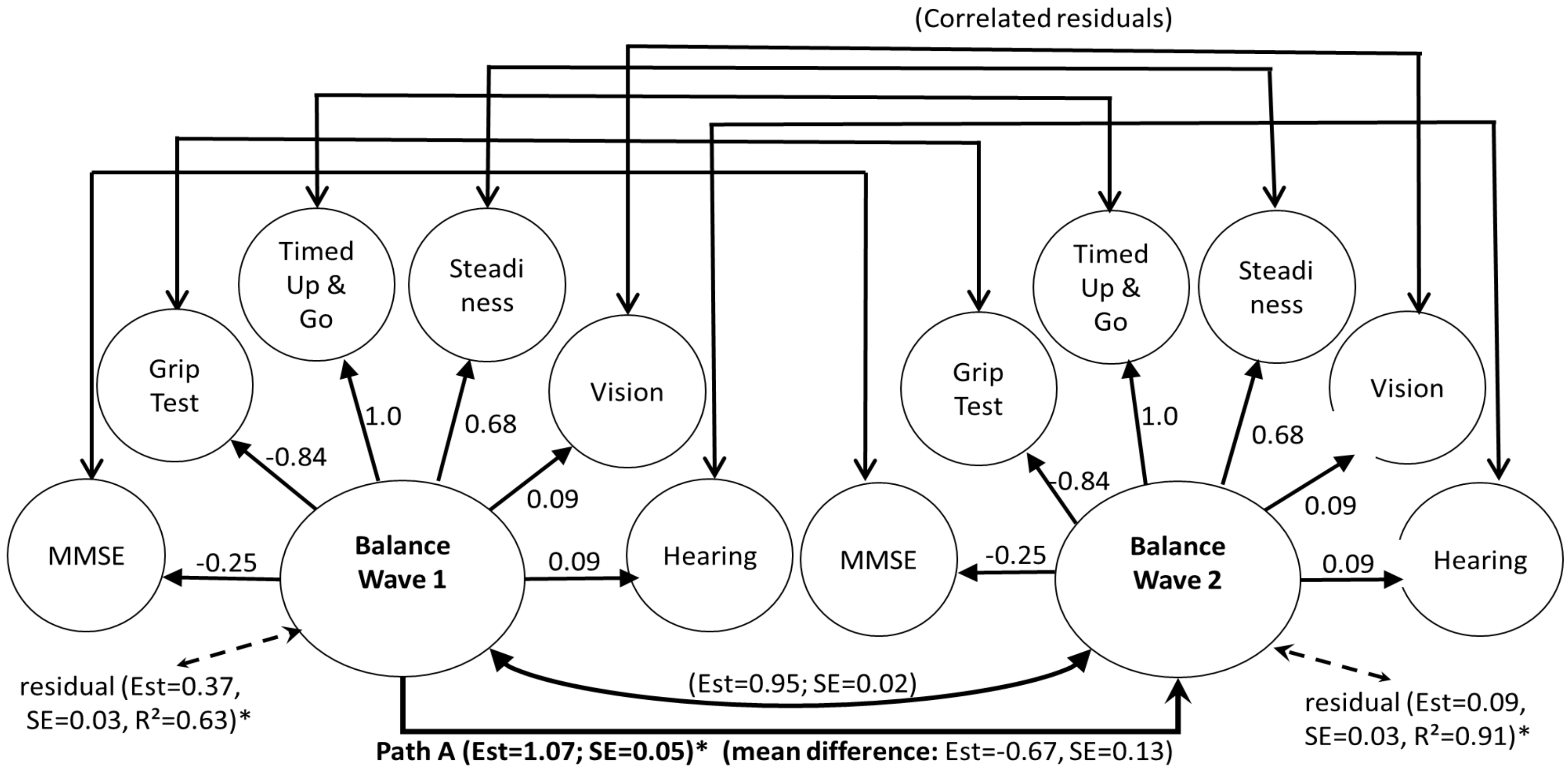
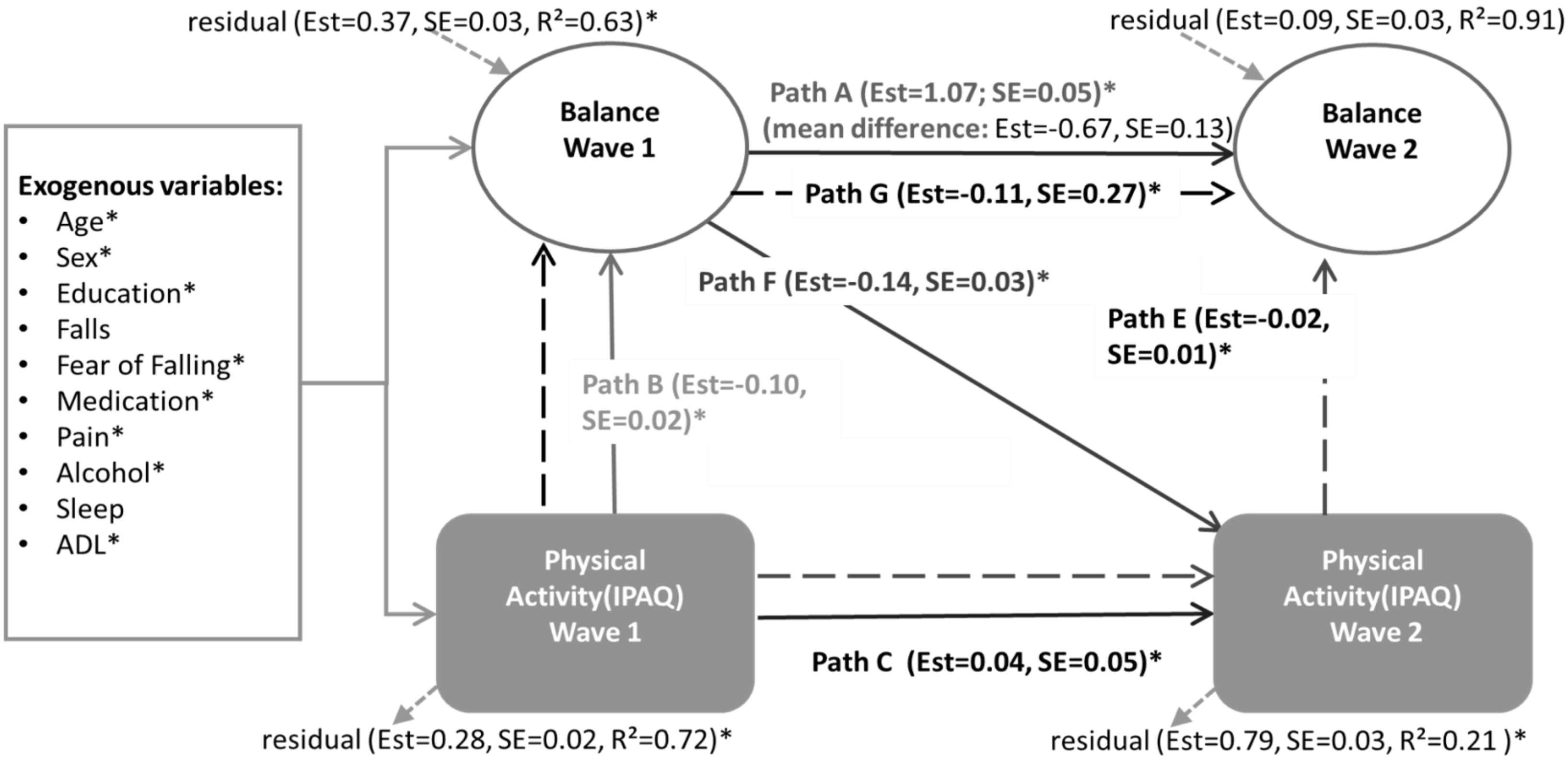


Figure 3.2
Results from confirmatory factor analysis of the composite measure of balance using multiple observed measures from TILDA



Note: Solid lines indicate direct effect and dashed line indicates indirect effect

Figure 3.3
Results from exploring the relationship between free-living PA and balance over a 2-year period controlling for covariates

Chapter Four:

Developing the composite measure of Physical Activity (PA) from the English Longitudinal Study of Ageing (ELSA)

4.1 Abstract

Background

The English Longitudinal Study of Ageing (ELSA) provides a categorical measure of PA which includes information relating to PA type and frequency that need to be combined for the purposes of analysis. Recommended methods of developing composite measures from categorical measures include either a conventional scoring approach that arbitrarily assigns individuals to groups based on specific characteristics, or a latent class analysis (LCA) approach which assigns individuals to groups based on the probability of being in that group. However, guidance is lacking regarding which approach is best more appropriate.

Methods

Three different composite measures of PA were calculated based on the three categorical questions from the ELSA study which provided information relating to type and frequency of PA. Firstly, a series of LCA models specifying subgroups from two to five were created, and using fit statistics (e.g. Bayesian Information Criterion; Akaike Information Criterion; Lo-Mendell-Ruben test; and entropy), a model of three subgroups was selected (inactive; LPA; and MVPA). An additional two composite measures were calculated using conventional approaches previously used to analyse the ELSA PA measure which included three subgroups (inactive, LPA, MVPA) (Demakakos et al., 2010) and two subgroups (inactive and MVPA) (Hamer et al., 2014).

To compare the three composite measures of PA a one-way analysis of variance (ANOVA) was carried out using indirect functional measures of balance from the ELSA study. Then, having identified the mean of PA and balance measures using each composite approach, the mean difference between the inactive and highest level of PA intensity subgroups was calculated. A higher mean difference was considered a greater level of distinction between subgroups and therefore a better approach. To identify where approaches diverged from agreement on classification of subgroups cross-tabulations and inspection of the raw data were explored to identify the level of misclassification.

Results

Descriptive statistics showed that 5,958 participants responded to the three PA questions, and there were 63 response patterns with complete data. Model fit indices identified that a class of three (inactive, LPA, and MVPA) described the data. A comparison of the means across all measures of balance (15 measures) showed that the level of distinction between the means was greater for the LCA (adjusted for measurement error) compared with the Hamer et al. (2014) approach, and greater for eleven out of the fifteen measures compared with the Demakakos et al. (2010) approach. Cross-tabulation between approaches showed a high degree of agreement in inactive groups with 97% between LCA and Demakakos et al. (2010) approaches, and 100% between LCA and Hamer et al. (2014) approaches, but there were high levels of disagreement in LPA and

MVPA subgroups. For example, of the 2,717 individuals classified in the low intensity activity group by LCA, Demakakos et al. (2010) agreed on 52% whilst Hamer et al., (2014) did not include a LPA group. Of the 6,631 individuals classified as MVPA groups by the LCA approach, Demakakos et al. (2010) agreed on 38% and Hamer et al. (2014) agreed on 38%. Further inspection of the overall effect of misclassification confirmed that the greatest discrepancies (2,722 participants or 46%) lay in the classification of MVPA where despite taking part in low or moderate activity, participants were classified as inactive by traditional scoring approaches.

Conclusion

This study shows that a latent class solution with three subgroups is the most appropriate approach to calculating the composite measure of PA intensity using the ELSA PA measure. It provides a greater distinction between groups than conventional scoring approaches, and highlights that an LCA approach provides a flexible approach that identifies patterns of activity in complex data that are ignored by conventional scoring approaches. Therefore, an LCA approach may allow more tailored interventions to be developed that maximise the benefits for an older adult population (≥50 years).

4.2 Introduction

PA data are often simplified into categorical classifications that need to be summarised into a composite measure for the purposes of analysis (Maslovskaya et al., 2018). However, summarising information can bias results due to overlapping of the distribution of the composite measure and measurement error (Jacobs, Goddard & Smith, 2005). Conventional scoring and latent class analysis (LCA) are two methods for summarising categorical data (Lanza & Rhoades, 2013; Maslovskaya et al., 2018; Nylund et al, 2007), but guidance on the most appropriate method is lacking (Maslovskaya et al., 2018).

For example, ELSA collects a categorical self-reported measure of PA at six different timepoints over a 10-year period using three categorical questions. Participants were asked to recall the frequency over the past year (more than once per week, once per week, one to three times per month, or hardly ever) of carrying out types of LPA such as walking for leisure, light housework, light gardening; MPA such as brisk walking, dancing, gardening, sports; and VPA such as running, fast cycling, aerobics, competitive sport (Ainsworth et al., 2011; WHO, 2018). Participants could potentially respond to the three questions relating to PA in 64 (4^3) different ways or patterns. Therefore, reducing the data (64 patterns of responses) into a composite measure, whilst minimising loss of information relating to response patterns and deriving the most parsimonious model is challenging.

Conventional scoring approaches include summarised scores of correct answers, true/false scales, Likert scales, or combinations of all these approaches (Maslovskaya et al., 2018). Demakakos et al. (2010) use a conventional scoring approach to explore the association between PA and diabetes using the ELSA PA variable by firstly dichotomising each question around the frequency cut-off point of ‘at least once per week or more’ within each category, and then using a summed index of the responses to derive three new variables relating to PA intensity: (1) Physically inactive, (2) LPA but not MVPA intensity at least once per week, and (3) MVPA at least once per week. Additionally, Hamer et al. (2014) explored the association between PA, risk of depression, and inflammatory mediators, by creating a binary variable from Demakakos et al.’s (2010) summed index, resulting in two new measures of PA level: (1) inactive or (2) MVPA group. Alternatively, an LCA approach, a subset of Structural Equation Modelling (SEM), proposes that there is an underlying unobserved construct of PA intensity that divides participants into mutually exclusive and exhaustive classes or subgroups (Lanza & Rhodes, 2013). Class or subgroup membership is inferred using actual data relating to the pattern of responses to the three questions for both type and frequency of PA rather than arbitrary cut-off points (Kongstead & Nielsen, 2017; Lanza & Rhoades, 2013; Nylund et al, 2007). Additionally, LCA uses multiple classes to estimate the unobserved construct of PA and as a result separates variance that is common among the observed measures that relate to the unobserved variable of PA from variance due to other factors such as those relating to bias due to, for example, the use of self-reported measures (Bauman et al., 2006; Dyrstad et al., 2014; Murphy, 2009; Saelens et al., 2012). As a result, LCA addresses measurement error (McCoach et al., 2007). LCA has been previously used in studies relating to PA where for example Jiang et al. (2016) (n= 1,344 participants) explored stages of change for regular exercise interventions in relation to diabetes prevention; Mooney et al. (2015) (n= 3,497 participants) identified individual and neighbourhood characteristics associated with patterns of PA in older adults; and Silverwood et al. (2011) (n= 3,847 participants) used LCA to characterise patterns of PA in adults aged 31-53 years old. An LCA approach has not been used to date to explore the PA data from the ELSA study. Therefore, this study will evaluate conventional scoring and LCA approaches to identify which approach is the most appropriate for summarising the PA measure provided by ELSA.

4.3 Aims and objectives

The overall aim of this study is to calculate the most appropriate composite measure of PA intensity using the ELSA data, and the following objectives were identified:

- To develop the composite measure of PA intensity using an LCA approach and identify the most parsimonious number of classes that describe the data.

- To calculate the composite measure of PA using conventional scoring approaches used by Demakakos et al. (2010), and Hamer et al. (2014).
- To evaluate the most appropriate approach by comparing the level of distinction made within groups using measures of balance from the ELSA study, where a greater level of distinction identifies a better approach (Hirji & Fagerland, 2009; Jacobs et al., 2005; Murphy, 2018).

4.4. Methods

4.4.1 Study design

The data used in this study were collected from ELSA, a nationally representative study of men and woman ≥50 years. The data were provided free of charge and access was achieved through an online registration process (http://ukdataservice.ac.uk/get-data/how-to-access/conditions.aspx.).

Key strengths of the ELSA study are that it includes repeated measures of the same individuals across multiple waves (six waves) over a 10-year period, thus facilitating longitudinal analyses of health outcomes; it includes multiple measures of health, physical and cognitive performance; and it is harmonised with other national studies of ageing such as TILDA (Appendix VIII) through its adherence to the Gateway to Global Ageing Initiative (GGAI) (https://g2aging.org/) therefore facilitating nationwide comparisons.

Sample: Participants were eligible to take part in ELSA if they had participated in the Health Survey for England (HSE) in 1998, 1999 and 2001 and agreed to follow-up; were born before 1 March 1952; and lived in a private household in England at baseline (2002-2003). The individual response rate was 67% and the total sample of 12,099 consisted of 11,391 core members, 636 partners aged > 50 years and 72 new partners aged 55 years. The mean age of the core sample was 65 years (range 50 to 100). The sample was refreshed at wave three to maintain the representation of people aged 50–53 years. A further refreshment sample of individuals aged 50–75 years was added at wave four, and refreshment of people aged 50–55 years was added at wave six. Apart from the range for year of birth, the eligibility criteria for refreshment samples remained the same as those for wave one. Repeated measures were recorded in the same individuals where 82% of wave one respondents participated in wave two, 73% in wave three, 74% in wave four, and 78% in wave five (Steptoe et al., 2013).

Participants were interviewed at two-yearly intervals using a Computer Assisted Personal Interview (CAPI), completed cognitive function tests and a walking speed test, as well as a Self-Completion Questionnaire (SCQ). On alternate waves (waves two, four and six), a nurse visit was carried out to complete a Health Assessment (HA) consisting of the collection of biomarkers and more detailed measures of function.

ELSA participants provided written informed consent, and the London Multi-Centre Research Ethics Committee granted ethical approval.

4.4.2 Measures

The appropriate PA, and balance variables were identified in the raw ELSA data files using a combination of the user guides and derived variables guides for each wave of ELSA data across six timepoints. These were then prepared for analysis in M-plus (version 7.4) in accordance with the M-plus user manual (Muthén & Muthén, 1998-2017) as detailed in Table 4.1 (page lxv). Measures were selected for this study that were similar to the measures used to develop the model of free-living physical activity and balance using the TILDA data (Appendix VIII).

4.4.2.1 Physical activity

ELSA asks participants three questions relating to type and frequency of PA. The questions were 1) How often in the last twelve months have you participated in vigorous activity? 2) How often in the last twelve months have you participated in moderate activity? and 3) How often in the last twelve months have you participated in mild activity? The options were: (a) more than once per week, (b) once per week, (c) one to three times per month, or (d) hardly ever. To assist in answering the questions, prompt cards with examples of activities categorised by intensity were shown where examples of mild physical activity included vacuuming, home repairs and laundry; examples of moderate physical activity included washing the car, dancing, floor/stretching exercises, walking at a moderate pace and gardening; and examples of vigorous physical activity included running or jogging, cycling, tennis, swimming and digging with a spade.

4.4.2.2 Balance measures

Sensory System

Self-rated vision was assessed using three questions: 1) How is your eyesight? 2) How is your eyesight seeing at a distance? and 3) How is your eyesight seeing close-up? where 5 is excellent; 4 is very good; 3 is good; 2 is fair; and 1 is poor. A high score indicates good balance.

Self-rated hearing was assessed using two questions: 1) How is your hearing? where 5 is excellent; 4 is very good; 3 is good; 2 is fair; and 1 is poor; and 2) Do you find it difficult to follow a conversation with background noise? Yes or no. A high score indicates good balance.

Cognitive system

Cognitive ability was assessed using five questions: 1) orientation in time was assessed by asking participants to recall the date; 2) prospective memory was assessed by asking participants to remember to recall ten words at a delayed timeframe; 3) verbal learning was assessed by asking participants to immediately recall ten words that were presented orally via a computer-based recording; 4) verbal fluency was assessed by asking participants to name as many different animals as possible in one minute; and 5) prospective memory was assessed by asking participants to remember to write their initials at a predefined time. A high score indicates good cognitive function.

Neuromuscular system

Hand grip strength (kg) was assessed using three measures of the dominant hand (Kg) using a hand-held dynamometer. A high score indicates good strength.

Lower body strength was assessed using one measure of the chair stand test, where the time taken to rise from a chair to a full standing position five times with arms folded across the chest was recorded (secs). The test incorporated the use of respondent's own armless, straight backed chair. A low time indicates good balance.

Gait speed was assessed only in participants ≥60 years by measuring the time taken to walk eight feet at usual pace (secs). Two attempts of the test of gait speed were included. A low time indicates good balance.

Steadiness was assessed using two questions: (1) How often do you have problems keeping balance when walking? Yes or no; (2) How often do you have problems with dizziness when walking? 1) never walks; 2) always; 3) very often; 4) sometimes; 5) hardly ever and 6) never.

4.5 Statistical analysis

In the current study a robust form of Maximum likelihood Estimation (MLE) was used to correct for non-normality. Missing data were assumed to be missing at random (Schafer & Graham, 2002), and a robust MLE approach uses a model-based estimation strategy to address missing data (Enders, 2013; Yaun & Bentler, 2000).

4.5.1 Conventional scoring approach

Two conventional scoring approaches are used for comparison in this study. Firstly, using the Demakakos et al. (2010) conventional score approach, the ELSA PA variable was dichotomised around the frequency cut off point of ‘at least once per week or more’ within each category and then a summed index of the responses was derived resulting in three new variables (1) Physically inactive; (2) LPA but not MVPA at least once per week; and (3) MVPA at least once per week. Then, using the Demakakos et al. (2010) summed index, the approach adopted by Hamer et al. (2014) was used to create a second composite score by creating a binary variable, resulting in an active or MVPA group. Both conventional scoring approaches assume data are missing at random.

4.5.2 Latent class analysis (LCA) approach

Using the ELSA PA, variable latent class analysis (LCA) models were successively fitted starting with a two-subgroup model and adding another subgroup for each successive model fitted in m-plus (Table 4.2, page lxxv). Model fit was assessed using model based measures of fit such as the Bayesian Information Criterion (BIC; Schwartz, 1978; Nylund et al., 2007), and the Akaike Information Criterion (AIC; Akaike, 1987; Kongstead & Nielsen, 2017), where a decrease in either indicates better model fit (Nylund et al., 2007; Kongstead & Nielsen, 2017); the Lo-Mendell-Ruben test (Lo et al., 2001; Nylund et al., 2007), where the p value indicates whether each model should be rejected in favour of another model; and‘ entropy’ a measure of classification distinction between subgroups was also used, where values near one indicate a high certainty in classification (Nylund et al., 2007). The strategy adopted was to identify a simple but distinct classification so if statistical fit indices suggested complex models with diminutive improvement, and models with fewer subgroups gave similar levels of subgroup distinction, then the model with fewer subgroups was selected.

A ‘BCH’ (Bolck et al., 2004; Bakk & Vermunt, 2015) approach as advocated by Asparouhov and Muthén (2014) was used as it is recommended for use with continuous distal outcomes and addresses both measurement error and classification error (Asparouhov & Muthen, 2014). The final model for the LCA of three classes is shown in Appendix XII.

4.5.3 Comparisons between the conventional score and LCA approaches

Firstly, the mean difference between each approach (conventional scoring and LCA) and the variables of balance from the ELSA study which included the sensory measures of self-rated vision and hearing; neuromuscular measures of self-rated steadiness, grip test, gait speed, and chair rise test; and cognitive measures of orientation in time, prospective memory (delayed and immediate), verbal learning, and verbal fluency was calculated using a one-way analysis of variance (ANOVA) (SPSS IBM version 23). A greater distinction is represented by a larger mean difference score in the balance measures between the inactive and MVPA groups.

A cross tabulation was also run for the LCA and each conventional scoring approach to explore the extent of misclassification between the approaches and the different groups. Finally, an inspection of the raw data was carried out to understand the nature of any divergence identified in the classification of groups and the extent of misclassification.

4.6 Results

Table 4.3 (page lxxvi) shows that there were 63 actual response patterns within the PA with complete data (an additional three response patterns were identified with incomplete data).

Table 4.2 (page lxxv) shows the model fit indices of each LCA of physical activity. A latent class of three was considered the best fit as the Lo-Mendell-Rubin Adjusted (LRT) test was significant for a latent class of three, but not for a latent class of four; there was a larger decrease in AIC, BIC and sample adjusted BIC values between classes two and three (1%, 1%, and 1% respectively) than classes three and four (0.1%, 0.1%, and 0.02% respectively); Pearson's chi-squared value was significant for three rather than four classes; and entropy for class three was good (0.8).

Using a one-way ANOVA descriptive statistic, the mean of each composite measure of PA on the observed measures of balance was carried out. The mean difference between the inactive and MVPA groups for each approach was calculated. A comparison of the mean differences using each approach was then carried out. Table 4.4 (page lxxvii) shows the mean and standard error (SE) obtained for each measure of balance along with the mean difference percentage shown between groups. The results show that the level of distinction between the means are greater for the LCA (adjusted for measurement error) and the Demakakos et al. (2010) approach across all measures compared with the Hamer et al. (2014) approach. Additionally, most measures for balance (eleven out of the fifteen) showed a greater distinction for the LCA approach compared with the Demakakos et al. (2010) approach, with one measure having the same mean difference (Eara1), and three measures having a lower mean difference using the LCA approach compared with the Demakakos et al. (2010) approach (Earb1, M2C, cfmd1).

Furthermore, cross tabulation exploring each of the approaches (Table 4.5, page lxxxii and Table 4.6, page lxxxiii) shows that within the inactive groups there was a high degree of agreement where LCA and Demakakos et al. (2010) approaches agreed on 97% and LCA and Hamer et al. (2014) approaches agreed on 100% of classifications. Of the 2,717 individuals classified as low intensity activity groups by LCA, 52% are classified as low, but 32% are classified as inactive and 16% as mod/vig intensity activity by Demakakos et al. (2010). Additionally, within the low intensity group, Hamer et al. (2014) classified 16% as MVPA and 84% as inactive. Of the 6, 631 individuals classified as MVPA groups by the LCA approach, there is only a 38% agreement in classification by Demakakos et al. (2010), whilst 60% are classified as low intensity, and 2% as inactive by Demakakos et al. (2010). Furthermore, within the MVPA group, Hamer et al. (2014) classified 38% as MVPA but 62% as Inactive.

The nature of the high level of divergence across groups was then explored. Table 4.7 (page lxxxiv) shows the responses to the original three PA questions, the corresponding classification derived from each of the approaches (LCA, Demakakos et al., 2010, Hamer et al., 2014), and the number affected by the misclassification. Table 4.8 (page lxxxvii) identifies the summary of patterns of misclassification. The patterns of misclassification confirmed that the greatest discrepancies lie in the classification of MVPA groups. For example, 1,990 (33%) participants classified in the MVPA group by LCA and LPA group by Demakakos et al. (2010), were classified as inactive by Hamer et al. (2014). Yet, these participants carried out mild PA more than once per week, as well as low PA at least one to three times per month (except for 34 participants who hardly ever carried out low PA). Additionally, 732 (12%) participants were classified by LCA and Demakakos et al. (2010) in the LPA group, but by Hamer et al. (2014) in the inactive group, but these participants carried out both vigorous and mild PA at least one to three times per month. Furthermore, 430 (7%) participants were classified by LCA in the LPA group, but by Demakakos et al. (2010) and Hamer et al. (2014) in the inactive group, but the majority (419) took part in mild PA at least one to three times per month. Also, 236 (4%) participants were classified in the LPA group by LCA but were classified in the MVPA group by both Demakakos et al. (2010) and Hamer et al. (2014). Yet, all these participants took part in vigorous, mild and low PA at least once per week.

4.7 Discussion

4.7.1 Principal findings

The fit statistics found that a latent class of three groups was the best solution to fit the 63 optimal patterns of responses to the physical activity (PA) questions. The mean difference analysis showed that the LCA showed a greater distinction between groups than traditional scoring approaches across most of balance measures (11/15, 73%). Additionally, cross tabulations between the methods showed agreement between the approaches on inactive groups, some agreement on LPA groups, but a high degree of divergence on MVPA groups. The nature of the high divergence on MVPA groups between the different approaches was explored, and misclassification was high with 63% of participants being misclassified. The data showed that whilst Hamer et al.'s (2014) approach provided the most parsimonious description of the data with only two groups (inactive and MVPA) the loss of information, as might be expected, is more significant compared with LCA and Demakakos et al. (2010) approaches. Furthermore, whilst Demakakos et al. (2010) provides a better description of the data than Hamer et al., (2010) there is still loss of information in the LPA and MVPA groups. For example, in both Demakakos et al., (2010) and Hamer et al.,

(2014) participants were classed as inactive despite carrying out LPA or MVPA less than once per week but more than one to three times per month. An LCA approach places a person in a group based on a probability so they are classified according to their weight across the groups, and therefore individuals are placed in a group with a given probability and as the mean differences results indicate. Thus, an LCA approach recognises that individuals who carry out PA that is LPA or MVPA once per week or one to three times per month are still active, thus providing a better description of the data. Therefore, an LCA approach is a more flexible approach that addresses individual differences and provides a better approach to calculating the composite measure of PA intensity using the PA data from the ELSA study.

4.7.2 Strengths and weaknesses

This study compares the results from three approaches using the same data to fully understand and evaluate the extent of the differences in classification and so provides a comprehensive approach to assessing the approaches. For example, conventional scoring approaches vary in the methods used to summarise the data, so comparisons made with a single method may not provide the same results as other methods used within the conventional scoring approaches. Therefore, including two conventional scoring approaches provides a good basis for comparisons across approaches.

Another strength of this study is that it includes both subjective and objective variables of balance to assess the mean differences between approaches to ensure the conclusions drawn are not influenced by bias associated with self-reporting measures (Dyrstad et al., 2014; Murphy, 2009; Saelens et al., 2012). A key weakness of the LCA approach is that there is no gold standard available to decide on the best model fit, and defining the classes obtained is a subjective process and therefore open to bias (Lanza & Rhoades, 2013). However, unlike conventional scoring approaches, an LCA approach allows us to test how many groups best describe the data using fit statistics rather than arbitrary group assignment based on characteristics that might not be the same for every individual (Bartholomew et al., 2008). Also, an LCA approach minimises measurement error, and uses the actual data to develop the classes without a significant loss of information (Bartholomew et al., 2008). Thus, an LCA approach makes the differences between groups easier to detect, allowing a potentially better prediction of treatment responses or more tailored and effective interventions to be designed (Lanza & Rhoades, 2013; Maslovskaya et al., 2018).

4.7.3 Future research and clinical implications

Future research should adopt an LCA approach involving the analysis of PA to understand the association with balance over time using the ELSA data.

4.8 Conclusion

A latent class approach makes a greater distinction between groups and helps to tailor interventions and maximise effects. The use of the LCA PA measure developed in this study will help to improve the robustness of the findings from the ELSA analyses carried out in Chapter five.

Table 4.1
Table showing balance variables used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						NOTES
			Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	
Heeye	How is your eyesight?	1=excellent; 2=very good; 3=good; 4=fair 5=poor Recoded into same variable so now: 5=excellent; 4=very good; 3=good; 2=fair 1=poor	Eyea1	Eyea2	Eyea3	Eyea4	Eyea5	Eyea6	Post recode high is good for balance
Hefrnd	How is your eyesight seeing at a distance?	1=excellent; 2=very good; 3=good; 4=fair 5=poor Recoded into same variable so now: 5=excellent; 4=very good; 3=good; 2=fair;1=poor	Eyeb1	Eyeb2	Eyeb3	Eyeb4	Eyeb5	Eyeb6	Post recode high is good for balance

Table 4.1 (continued)
Table showing balance variables used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						
			Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	NOTES
Hepap	How is your eyesight seeing close up?	1=excellent; 2=very good; 3=good; 4=fair 5=poor Recoded into same variable so now: 5=excellent; 4=very good; 3=good; 2=fair 1=poor	Eyec1	Eyec2	Eyec3	Eyec4	Eyec5	Eyec6	Post recode High is good for balance
Hehear	How is your hearing?	1=excellent; 2=very good; 3=good; 4=fair; 5=poor Recoded into same variable so now: 5=excellent; 4=very good; 3=good; 2=fair 1=poor	Eara1	Eara2	Eara3	Eara4	Eara5	Eara6	Post recode high is good for balance
Hehra	Do you find it difficult to follow a conversation with background noise?	Yes=1; No=2 Recoded so Yes=2; No=1	Earb_1 Original variable is Earb1	Earb_2 Earb2	Earb_3 Earb3	Earb_4 Earb4	Earb_5 Earb5	Earb_6 Earb6	Post recode high is good for balance

Table 4.1 (continued)
Table showing balance variables used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						NOTES
			Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	
Hebal	How often do you have problems keeping balance when walking?	1=always; 2=very often 3=often; 4=sometimes 5=never; 6=never walks	Steada 1	Steada 2	-	Steada 4	-	-	No measures for waves 3, 5, & 6 High is good for balance
Hediz	How often do you have problems with dizziness when walking?	1=always; 2=very often; 3=often; 4=sometimes; 5=never; 6=never walks	Steadb 1	Steadb 2	-	Steadb 4	-	-	No measures for waves 3, 5, & 6 High is good for balance
Mmwlka	Gait speed	Seconds Recoded -reverse scored and then centred	-	Gait_a 2	Gait_a 3	Gait_a 4	Gait_a 5	Gait_a 6	Wave 1 not include due to inconsistency in data-confirmed by ELSA project manager Post recode: High on gait then good on balance

Table 4.1 (continued)
Table showing balance variables used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						NOTES
			Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	
Mmwlhb	Gait speed	Seconds	-	Gait_b 2	Gait_b 3	Gait_b 4	Gait_b 5	Gait_b 6	Wave 1 not included due to inconsistency in recording the data-confirmed by ELSA project manager Post recode: High on gait then good on balance
cfDscr	Date recall questions (4 questions)	Maximum score of 4	Cfra1	Cfra2	-	Cfra4	Cfra5	Cfra6	No measure for wave 3 High is good for balance
Cflisten	Immediate word recall (no. of words)	Maximum score of 10	Cfrb1	Cfrb2	-	Cfrb4	Cfrb5	Cfrb6	No measure for wave 3 High is good for balance
cfAni	No. animals (fluency)	Number	Cfrc1	Cfrc2	-	Cfrc4	Cfrc5	Cfrc6	No measure for wave 3 High is good for balance

Table 4.1 (continued)
Table showing balance variables used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						NOTES
			Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	
Cffascr	Remembering to write initials (prospective memory score)	Maximum score 5	Cfmd1	Cfmd2	-	Cfmd4	Cfmd5	Cfmd6	No measure for wave3 High is good for balance
Cflisd	Delayed word recall	Maximum score 10	Cfre1	cfre2	-	Cfre4	Cfre5	Cfre6	No measure for wave3 High is good for balance
Mmgsd1	Grip strength measure 1	Kg	-	Gripa2	-	Gripa4	-	Gripa6	Nurse assessment waves only wave 3, 4, 6 High is good for balance
Mmgsd2	Grip strength measure 2	Kg	-	Gripb2	-	Gripb4	-	Gripb6	Nurse assessment waves only wave 3, 4, 6 High is good for balance

Table 4.1 (continued)
Table showing balance variables used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						NOTES
			Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	
Mmgsd3	Grip strength measure 3	Kg	-	Gripc2	-	Gripc4	-	Gripc6	Nurse assessment waves only wave 3, 4, 6 High is good for balance
Mmrfti	Chair raise	Seconds Recoded: reverse scored and then centred	-	M2c	-	M4c	-	M6c	Nurse assessment waves only wave 3, 4, 6 High is good for balance

Table 4.1 (continued)
Table showing physical activity variables used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						NOTES
			Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	
heacta	How often do you take part in vigorous sports or activities?	1=more than once per week; 2=once per week; 3=one to three times per month; 4=hardly ever or never	Pa1	Pa2	Pa3	Pa4	Pa5	Pa6	Latent class analysis variable used in analysis
heactb	How often do you take part in moderate sports or activities?	1=more than once per week; 2=once per week; 3=one to three times per month; 4=hardly ever or never	Pb1	Pb2	Pb3	Pb4	Pb5	Pb6	
heactc	How often do you take part in mild sports or activities?	1=more than once per week; 2=once per week; 3=one to three times per month; 4=hardly ever or never	Pc1	Pc2	Pc3	Pc4	Pc5	Pc6	

Table 4.1 (continued)
Table showing the covariates used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						NOTES
			Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	
dhager	age	Actual age	Agec_ 1-	Agec_ 2	Agec_ 3	Agec_ 4	Agec_ 5	Agec_ 6	Centred age variable used (dhager (x) minus average age of wave)
disex	sex	Male = 1 Female = 2	Sex	-	-	-	-	-	Sex at baseline used for analysis
helim	Does an illness (es) or disability (ies) limit your activities in any way?	Yes=1 No= 2	ADLa1	ADLa2	ADLa3	ADLa4	ADLa5	ADLa6	Baseline ADLa1 used for analysis
heflb	Have you fallen down in the last two years (for any reason)?	Yes=1 No=2	Falla1	Falla2	Falla3	Falla4	Falla5	Falla6	Baseline Falla1 used for analysis
hepain	Are you often troubled by pain?	Yes=1 No=2	Paina1	Paina2	Paina3	Paina4	Paina5	Paina6	Baseline Paina1 used for analysis

Table 4.1 (continued)
Table showing the covariates used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						NOTES
			Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	
edqual	Highest education level at wave (X)	1=NVQ4/NVQ5/Degree or equivalent; 2=higher education below degree; 3=NVQ3/GCE A level equivalent; 4= NVQ2/GCE O level equivalent; 5= NVQ1/CSE other grade equivalent; 6=foreign/other; 7=no qualification	EDU1	EDU2	EDU3	EDU4	EDU5	-	Baseline EDU1 used for analysis
scako	How often have you had an alcoholic drink during the last 12 months?	1=almost every day; 2=five or six days per week; 3=three or four days per week; 4=once or twice a week; 5=once or twice a month; 6=once every couple of months; 7=once or twice a year; 8=not at all in the last 12 months	-	Alca2	Alca3	Alca4	Alca5	Alca6	Baseline Alca2 used for analysis

Table 4.1 (continued)
Table showing the covariates used in the analysis of ELSA (original and recoded)

Original variable name	Description	Response	Recoded variable names						NOTES
			Wave 1	**Wave 2**	**Wave 3**	**Wave 4**	**Wave 5**	**Wave 6**	
heslpf	Sleep-rating quality overall	1=very good; 2=good; 3=fairly bad; 4=very bad	-	-	-	Sleepe 4	-	Sleepe 6	Baseline sleepe4 used for analysis
medcnjd	Are you taking any medication prescribed by a doctor/nurse ?	1=yes 2=no	-	-	-	-	-	MED	Baseline MED used for analysis

Table 4.2
Model fit indices of the Latent Class Analysis models of physical activity (PA) intensity – Inactive, Low, and Moderate/Vigorous groups

	Adjusted LRT Test		**Information Criteria**			**Chi squared**			**Entropy**
Number of subgroups	Value	P value	AIC	BIC	Sample size adjusted BIC	Value	df	P value	Classification quality
2	2096.90	0.0000	34078.91	34206.06	34145.69	5079.09	44	0.0000	0.8
3	396.96	0.0000	33697.38	33891.46	33799.30	108.54	34	0.0000	0.8
4	62.57	0.6253	33654.08	33915.09	33791.16	42.77	24	0.0106	0.8
5	24.03	0.9668	33649.78	33977.71	33821.99	16.76	14	0.2694	0.7

Note:

- Adjusted LRT is the adjusted Lo-Mendell-Rubin test (Lo et al., 2001; Nylund et al., 2007), where a non-significant result indicates the previous model should be accepted.
- AIC is the Akaike Information Criterion (Akaike, 1987; Kongstead and Nielsen, 2017), where decrease indicates a better fit.
- BIC is the Bayesian Information Criterion (Schwartz, 1978; Nylund et al., 2007), where decrease indicates better fit.
- Entropy values near one indicate a high certainty in classification (Nylund et al., 2007).

Table 4.3
Response patterns and observed frequencies for the physical activity questions from the ELSA dataset

1. 000(793)	2. 100(362)	3. 200(330)	4. 300(1432)	5. 010(66)	6. 110(103)
7. 210(80)	8. 310(463)	9. 020(11)	10. 120(15)	11. 220(21)	12. 320(204)
13. 030(18)	14. 130(13)	15. 230(7)	16. 330(417)	17. 001(56)	18. 101(27)
19. 201(24)	20. 301(121)	21. 011(24)	22. 111(39)	23. 211(29)	24. 311(112)
25. 021(4)	26. 121(1)	27. 221(13)	28. 321(60)	29. 031(7)	30. 131(4)
31. 231(2)	32. 331(167)	33. 002(20)	34. 102(7)	35. 202(7)	36. 302(45)
37. 112(6)	38. 212(8)	39. 312(28)	40. 022(6)	41. 122(1)	42. 222(10)
43. 322(31)	44. 032(1)	45. 132(2)	46. 232(2)	47. 332(54)	48. 003(23)
49. 103(11)	50. 203(6)	51. 303(74)	52. 013(4)	53. 113(5)	54. 213(3)
55. 313(30)	56. 023(4)	57. 123(1)	58. 223(1)	59. 323(18)	60. 033(10)
61. 133(7)	62. 233(4)	63. 333(501)	64. 3*0 (1)	65. 33*(1)	66. *00(1)

Note:

-pattern of responses shown with frequency of pattern shown in brackets

-* denotes missing response

- The questions were 1) How often in the last twelve months have you participated in vigorous activity? 2) How often in the last twelve months have you participated in moderate activity? and 3) How often in the last twelve months have you participated in mild activity. The options were: (0) more than once per week, (1) once per week, (2) one to three times per month, or (3) hardly ever.

Table 4.4
Means and mean difference of the balance variables using a latent class approach (LCA), and conventional scoring approaches by Demakakos et al. (2010) and Hamer et al. (2014)

Balance measures	**Latent class approach Mean (SE)** [mean not adjusted for measurement error]	**LCA mean difference between Inactive and Mod/Vig** [unadjusted for measurement error]	**Demakakos et al. (2010) Mean (SE)**	**Demakakos et al. (2010) mean difference between Inactive and Mod/Vig**	**Hamer et al. (2014) Mean (SE)**	**Hamer et al. (2014) mean difference between Inactive and Mod/Vig**
Gaita2 (lower score is good)	Inactive: 5.08 (0.15) [4.97] Low: 3.36 (0.06) [3.39] Mod/vig: 2.99 (0.03) [3.03]	41% [39%]	Inactive: 4.57 (0.10) Low: 3.24 (0.03) Mod/vig: 2.79 (0.03)	39%	Inactive: 3.70 (0.04) Mod/vig: 2.79 (0.03)	25%
Steada1 (higher score is good)	Inactive: 4.04 (0.04) [4.08] Low: 4.69 (0.02) [4.70] Mod/vig: 4.82 (0.01) [4.81]	19% [18%]	Inactive: 4.24 (0.02) Low: 4.75 (0.01) Mod/vig: 4.87 (0.01)	15%	Inactive: 4.56 (0.01) Mod/vig: 4.87 (0.01)	7%
Steadb1 (higher score is good)	Inactive: 4.55 (0.03) [4.57] Low: 4.83 (0.02) [4.84] Mod/vig: 4.89 (0.01) [4.89]	7% [7%]	Inactive: 4.64 (0.01) Low: 4.86 (0.01) Mod/vig: 4.92 (0.01)	6%	Inactive: 4.78 (0.01) Mod/vig: 4.92 (0.01)	3%

Table 4.4 (continued)
Means and mean difference of the balance variables using a latent class approach (LCA), and conventional scoring approaches by Demakakos et al. (2010) and Hamer et al. (2014)

Balance measures	**Latent class approach Mean (SE)** [mean not adjusted for measurement error]	**LCA mean difference between Inactive and Mod/Vig** [unadjusted for measurement error]	**Demakakos et al. (2010) Mean (SE)**	**Demakakos et al. (2010) mean difference between Inactive and Mod/Vig**	**Hamer et al. (2014) Mean (SE)**	**Hamer et al. (2014) mean difference between Inactive and Mod/Vig**
Eara1 (higher score is good)	Inactive: 3.19 (0.04) [3.23] Low: 3.43 (0.04) [3.47] Mod/vig: 3.58 (0.02) [3.57]	12% [11%]	Inactive: 3.28 (0.02) Low: 3.51 (0.02) Mod/vig: 3.66 (0.02)	12%	Inactive: 3.42 (0.01) mod/vig: 3.66 (0.02)	7%
Earb1 (higher score is good)	Inactive: 1.60 (0.02) [1.60] Low: 1.65 (0.02) [1.67] Mod/vig: 1.69 (0.01) [1.68]	6% [5%]	Inactive: 1.61 (0.01) Low: 1.67 (0.01) Mod/vig: 1.72 (0.01)	7%	Inactive: 1.65 (0.01) mod/vig: 1.72 (0.01)	4%
Eyea1 (higher score is good)	Inactive: 3.04 (0.03) [3.07] Low: 3.39 (0.03) [3.42] Mod/vig: 3.60 (0.02) [3.60]	18% [17%]	Inactive: 3.14 (0.02) Low: 3.52 (0.01) Mod/vig: 3.66 (0.02)	17%	Inactive: 3.37 (0.01) Mod/vig: 3.66 (0.02)	9%

Table 4.4 (continued)
Means and mean difference of the balance variables using a latent class approach (LCA), and conventional scoring approaches by Demakakos et al. (2010) and Hamer et al. (2014)

Balance measures	**Latent class approach Mean (SE)** [mean not adjusted for measurement error]	**LCA mean difference between Inactive and Mod/Vig** [unadjusted for measurement error]	**Demakakos et al. (2010) Mean (SE)**	**Demakakos et al. (2010) mean difference between Inactive and Mod/Vig**	**Hamer et al. (2014) Mean (SE)**	**Hamer et al. (2014) mean difference between Inactive and Mod/Vig**
Eyeb1 (higher score is good)	Inactive: 3.37 (0.03) [3.39] Low: 3.70 (0.03) [3.73] Mod/vig: 3.92 (0.02) [3.90]	16% [15%]	Inactive: 3.20 (0.02) Low: 3.51 (0.01) Mod/vig: 3.65 (0.02)	14%	Inactive: 3.65 (0.02) Mod/vig: 3.40 (0.01)	7%
Eyec1 (higher score is good)	Inactive: 3.18 (0.03) [3.24] Low: 3.50 (0.03) [3.55] Mod/vig: 3.74 (0.02) [3.72]	18% [15%]	Inactive: 3.29 (0.02) Low: 3.65 (0.01) Mod/vig: 3.79 (0.02)	15%	Inactive: 3.51 (0.01) Mod/vig: 3.79 (0.02)	8%
M2c (lower score is good)	Inactive: 5.02 (0.56) [4.83] Low: 4.26 (0.29) [4.02] Mod/Vig: 3.35 (0.13) [3.26]	33% [33%]	Inactive: 4.64 (0.19) Low: 3.65 (0.11) Mod/vig: 2.84 (0.12)	39%	Inactive: 3.98 (0.10) Mod/vig: 2.84 (0.12)	32%

Table 4.4 (continued)
Means and mean difference of the balance variables using a latent class approach (LCA), and conventional scoring approaches by Demakakos et al. (2010) and Hamer et al. (2014)

Balance measures	**Latent class approach Mean (SE)** [mean not adjusted for measurement error]	**LCA mean difference between Inactive and Mod/Vig** [unadjusted for measurement error]	**Demakakos et al. (2010) Mean (SE)**	**Demakakos et al. (2010) mean difference between Inactive and Mod/Vig**	**Hamer et al. (2014) Mean (SE)**	**Hamer et al. (2014) mean difference between Inactive and Mod/Vig**
Gripa2 (higher score is good)	Inactive: 22.99 (0.49) [23.88] Low: 28.72 (0.45) [28.71] Mod/vig: 30.29 (0.25) [30.39]	32% [27%]	Inactive: 25.68 (0.28) Low: 28.87 (0.19) Mod/vig: 31.79 (0.24)	24%	Inactive: 27.79 (0.24) Mod/vig: 31.79 (0.24)	14%
Cfra1 (higher score is good)	Inactive: 3.61 (0.02) [3.62] Low: 3.75 (0.02) [3.76] Mod/vig: 3.79 (0.01) [3.80]	5% [5%]	Inactive: 3.65 (0.01) Low: 3.78 (0.01) Mod/vig: 3.81 (0.01)	4%	Inactive: 3.73 (0.01) Mod/vig: 3.81 (0.01)	2%
Cfrb1 (higher score is good)	Inactive: 4.69 (0.06) [4.78] Low: 5.49 (0.05) [5.53] Mod/vig: 5.76 (0.03) [5.76]	23% [21%]	Inactive: 4.93 (0.03) Low: 5.62 (0.02) Mod/vig: 5.93 (0.03)	20%	Inactive: 5.36 (0.02) Mod/vig: 5.93 (0.29)	11%

Table 4.4 (continued)
Means and mean difference of the balance variables using a latent class approach (LCA), and conventional scoring approaches by Demakakos et al. (2010) and Hamer et al. (2014)

Balance measures	**Latent class approach Mean (SE)** [mean not adjusted for measurement error]	**LCA mean difference between Inactive and Mod/Vig** [unadjusted for measurement error]	**Demakakos et al. (2010) Mean (SE)**	**Demakakos et al. (2010) mean difference between Inactive and Mod/Vig**	**Hamer et al. (2014) Mean (SE)**	**Hamer et al. (2014) mean difference between Inactive and Mod/Vig**
Cffc1 (higher score is good)	Inactive: 16.25 (0.19) [16.27] Low: 19.11 (0.19) [19.16] Mod/vig: 20.59 (0.11) [20.54]	27% [26%]	Inactive: 17.14 (0.11) Low: 19.84 (0.08) Mod/vig: 21.15 (0.12)	23%	Inactive: 18.81 (0.07) Mod/vig: 21.15 (0.12)	12%
Cfmd1 (higher score is good)	Inactive: 3.83 (0.05) [3.64] Low: 3.93 (0.03) [3.83] Mod/vig: 3.59 (0.05) [3.93]	6% [8%]	Inactive: 3.67 (0.03) Low: 3.88 (0.02) Mod/vig: 3.99 (0.03)	9%	Inactive: 3.80 (0.02) Mod/vig: 3.99 (0.03)	5%
Cfre1 (higher score is good)	Inactive: 3.75 (0.06) [3.84] Low: 4.38 (0.06) [4.40] Mod/vig: 4.62 (0.03) [4.62]	23% [20%]	Inactive: 3.92 (0.03) Low: 4.53 (0.03) Mod/vig: 3.73 (0.03)	5%	Inactive: 4.31 (0.02) Mod/vig: 4.73 (0.03)	10%

***note**

- cfmd1 and cfre1 were checked as the order appeared to be inconsistent, but there are no spurious data in the dataset.
- mean and standard error (SE) shown in brackets.
- [ ] denotes mean post BHC treatment in Mplus to remove measurement error.

Table 4.5
Cross tabulation results between latent class analysis (LCA), and the conventional scoring approach used by Demakakos et al. (2010)

Demakakos et al., (2010) approach (DEM)	Latent class (3) LCA3			
	inactive	low intensity	Mod/vig. intensity	Total
Inactive	2485 (97%)	859 (32%)	129 (2%)	3473
Low intensity	73 (3%)	1424 (52%)	3994 (60%)	5491
Mod/vig. intensity	0 (0%)	434 (16%)	2508 (38%)	2942
Total	2558	2717	6631	11906

Table 4.6
Table showing the cross tabulation between latent class analysis (LCA), and the conventional scoring approach used by Hamer et al., (2014)

Hamer et al., (2014) approach (HAM)	Latent class (3) LCA3			Total
	inactive	low intensity	Mod/vig. intensity	
Inactive	2558 (100%)	2283 (84%)	4123 (62%)	8964
Mod/vig. intensity	0 (0%)	434 (16%)	2508 (38%)	2942
Total	2558	2717	6631	11906

Table 4.7

Results of the analysis of the extent of misclassification by latent class analysis (LCA), and conventional scoring approaches by Demakakos et al. (2010), and Hamer et al. (2014) using the original ELSA physical activity (PA) scoring (vigorous, mild, and low PA) (n=5,958 participants)

Original PA type scoring pattern (vigorous, mild, low) (corrected where *1=more than once per week; 2=once per week; 3= one to three times per month; 4=hardly ever or never*	**LCA PA intensity** *(1=inactive; 2=low; 3=mod/vig)*	**Demakakos et al. (2010) score PA intensity** *(1=inactive; 2=low; 3=mod/vig)*	**Hamer et al. (2014) score PA intensity** *(1=inactive; 3=mod/vig)*	**Total number of responses**	**Agreement between approaches (yes/no)**
322	2	2	1	29 (0.5%)	No
422	2	2	1	112 (1.9%)	No
132	2	2	1	4 (0.07%)	No
232	2	2	1	1 (0.02%)	No
331	2	1	1	13 (0.2%)	No
432	2	1	1	60 (1%)	No
143	1	1	1	7 (0.1%)	Yes inactive
242	1	2	1	4 (0.07%)	No
342	2	1	1	2 (0.03%)	No
442	1	1	1	167 (2.8%)	Yes inactive
113	3	2	1	20 (0.33%)	No
213	3	2	1	7 (0.1%)	No
313	3	1	1	7 (0.1%)	No
412	3	2	1	45 (0.76%)	No
223	2	2	1	6 (0.1%)	No
323	2	1	1	8 (0.13%)	No
433	2	1	1	31 (0.52%)	No
423	2	1	1	28 (0.47%)	No
133	2	1	1	6 (0.1%)	No
233	2	1	1	1 (0.02%)	No

Table 4.7 (continued)
Results of the analysis of the extent of misclassification by latent class analysis (LCA), and conventional scoring approaches by Demakakos et al. (2010), and Hamer et al. (2014) using the original ELSA physical activity (PA) scoring (vigorous, mild, and low PA) (n=5,958 participants)

Original PA type scoring pattern (vigorous, mild, low) (corrected where *1=more than once per week; 2=once per week; 3= one to three times per month; 4=hardly ever or never*	**LCA PA intensity** *(1=inactive; 2=low; 3=mod/vig)*	**Demakakos et al. (2010) score PA intensity** *(1=inactive; 2=low; 3=mod/vig)*	**Hamer et al. (2014) score PA intensity** *(1=inactive; 3=mod/vig)*	**Total number of responses**	**Agreement between approaches (yes/no)**
333	2	1	1	10 (0.17%)	No
143	1	1	1	1 (0.02%)	Yes inactive
243	1	1	1	2 (0.03%)	Yes inactive
343	2	1	1	2 (0.03%)	No
443	1	1	1	54 (0.9%)	Yes inactive
114	3	2	1	23 (0.39%)	No
214	3	2	1	11 (0.18%)	No
314	3	1	1	6 (0.1%)	No
414	1	1	1	74 (1.24%)	Yes inactive
444	1	1	1	501 (8.41%)	Yes inactive
424	2	1	1	30 (0.5%)	No
434	1	1	1	18 (0.3%)	Yes inactive
124	2	2	1	4 (0.07%)	No
224	2	2	1	5 (0.08%)	No

Table 4.7 (continued)
Results of the analysis of the extent of misclassification by latent class analysis (LCA), and conventional scoring approaches by Demakakos et al. (2010), and Hamer et al. (2014) using the original ELSA physical activity (PA) scoring (vigorous, mild, and low PA) (n=5,958 participants)

Original PA type scoring pattern (vigorous, mild, low) (corrected where *1=more than once per week; 2=once per week; 3= one to three times per month; 4=hardly ever or never*	**LCA PA intensity** *(1=inactive; 2=low; 3=mod/vig)*	**Demakakos et al. (2010) score PA intensity** *(1=inactive; 2=low; 3=mod/vig)*	**Hamer et al. (2014) score PA intensity** *(1=inactive; 3=mod/vig)*	**Total number of responses**	**Agreement between approaches (yes/no)**
324	2	1	1	**3 (0.05%)**	No
134	2	1	1	**4 (0.07%)**	No
234	2	1	1	**1 (0.02%)**	No
334	2	1	1	**1 (0.02%)**	No
144	1	1	1	10 (0.17%)	Yes inactive
244	1	1	1	7 (0.12%)	Yes inactive
4*1	3	1	1	**1 (0.02%)**	No
44*	1	1	1	1 (0.02%)	Yes inactive
*11	3	2	1	**1 (0.02%)**	No

**Note: ***Original scoring pattern from observed variables where three questions relating to type of activity were *vigorous, mild, low PA and frequency was 1=more than once per week; 2=once per week; 3= one to three times per month; 4=hardly ever or never. M-plus output of response pattern mapped onto actual response where 0=1; 1=2; 2=3; 3=4. For the purposes of comparison, the Hamer et al. (2014) scoring was changed from 0=mod/vig PA to 3=mod/vig.*

Table 4.8
Summary of misclassification patterns between the latent class analysis (LCA) and conventional score approaches (Demakakos et al., 2010; Hamer et al., 2014) and total population misclassified

LCA approach	Demakakos et al., (2010) approach	Hamer et al., (2014) approach	Number affected (response pattern shown in brackets)	Total number affected
Low	Mod/vig	Mod/vig	66 (121); 4 (122); 103 (221); 24 (122); 39 (222)	236 (4%)
Low	Low	Inactive	11 (131); 15 (231); 29 (322); 4 (132); 6 (223); 4 (124); 5 (224); 3 (324); 463 (421); 80 (321); 121 (412)	732 (12%)
Low	Inactive	inactive	204 (431); 10 (333); 2 (343); 30 ()424); 4 (134); 1 (234); 1 (334); 21 (331); 7 (341); 1 (232); 13 (331); 60 (432); 2 (342); 8 (323); 31 (433); 28 (423); 6 (133); 1 (233)	430 (7%)
Inactive	Low	Inactive	18 (141); 13 (241); 4 (242)	35 (0.6%)
Mod/vig	Low	Low	24 (312)	24 (0.4%)
Mod/vig	Low	Inactive	330 (311); 121 (411); 1432 (411); 20 (113); 7 (213); 45 (412); 23 (114); 11 (214); 1 (*11)	1,990 (33%)
Mod/vig	Inactive	Inactive	7 (313); 6 (314); 1 (4*1)	14 (0.2%)

Chapter Five:

Exploring the trajectory of change in balance performance associated with physical activity intensity in older adults (≥50 years) using the English Longitudinal Study of Ageing (ELSA)

5.1 Abstract

Background

Older adults are more likely to engage in low intensity physical activity (LPA), but perhaps because the focus of policy has been on MVPA, an understanding of the effects of LPA on balance is lacking. Further, research suggests that LPA requires a longer period of time for its benefits to be realised and studies exploring the effects of PA on balance have in the main been clinical trials or cross-sectional studies. As a result, an understanding of how PA affects balance over time is needed.

Methods

An understanding of individual differences and trajectory of change over time requires data collection over multiple timepoints (Bentein et al., 2005; Duncan & Duncan, 2009). Consequently, this study used data relating to PA and balance from the ELSA study that provided data across six time points over a 10-year period. Firstly, confirmatory factor analysis (CFA) was used to develop a robust composite measure of balance, and to identify the measures with the strongest relationship with balance. Then, using a PA measure developed using latent class analysis (LCA) that identified inactive, LPA, and MVPA groups, a latent growth model (LGM) was run using the three measures of gait speed, steadiness, and cognitive function (Cog). The covariates of age, sex, activities of daily living (ADL), pain, alcohol, history of falls, and education were introduced into the model where appropriate.

Results

The indirect measures of gait speed, cognitive function, and steadiness were identified as those measures with the strongest relationship with balance in older adults ≥50 years. Balance declined over a ten-year period by 29%. LGM analysis showed that the rate of decline across all measures at all ages was slower in the MVPA group than other PA groups. Additionally, exploring the effects of age showed that participants aged 60-70 years taking part in MVPA performed better on all three measures than other PA groups, but taking part in LPA was more beneficial than being inactive. Also, participants aged ≥70 years performed better on the measures of gait speed and cognitive function in the LPA group than other PA groups, whilst participants aged ≥70 years performed better on the measure of steadiness than other PA groups.

Conclusion

The neuromuscular measures of chair rise test, gait speed test, HT, and steadiness, the sensory measures of vision and hearing, and the cognitive measures identified from the ELSA study collectively provide a model for balance assessment in an older population in England. LGM analysis identified that higher intensity PA is better than lower intensity PA for older adults ≤70 years and can also reduce the decline in balance measures over time. However, LPA has more benefits for older adults aged ≥70 years for balance measures.

5.2 Introduction

As previously highlighted (Chapter one), healthcare promotion for fall prevention focuses on MVPA (CSP, 2018; DoH, 2011; NICE, 2013; PHE, 2017, 2018; RCP, 2015; WHO, 2007, 2010, 2015, 2018). However, evidence suggests that older adults are more likely to engage in LPA due to health-related issues or physical capability (Ainsworth et al., 2011; Arnardottir et al., 2013; Franco et al., 2015). However, whilst the benefits of LPA for general health are beginning to emerge (Demakakos et al., 2010; Dohrn et al., 2018; Tse et al., 2015), longitudinal evidence supporting the benefits for balance performance is lacking (Bauman et al., 2016; NICE, 2013; PAGAC, 2018; Schmidt et al., 2017) (Chapter one, section 1.2.2). Furthermore, whilst the findings from the TILDA analysis (Chapter three) showed that older adults (≥50 years) who are more active have better balance over a two-year period, only two time points were available (linear model), and so an understanding of individual differences and trajectory of change over time is limited (Bentein et al., 2005; Duncan & Duncan, 2009). Consequently, this study uses data relating to PA and balance from the ELSA study (described in Chapter four, section 4.4.1) that provides data across six time points over a 10-year period, and includes a measure of PA, and multiple measures of balance, to add to our understanding of the effects of different PA intensity on balance over time.

This study uses a Latent Growth Model (LGM) within a SEM approach to analyse the trajectory of change in PA and balance over time (Duncan & Duncan, 2009; Jung & Wickrama, 2008). This approach has specific benefits where for example model fit can be assessed using fit statistics; change can be assessed; and missing data as well as measurement error addressed (Beran et al., 2010). Furthermore, LGM contains both variable-centred and person-centred approaches, thus both change on balance performance at the ELSA population level, as well as at the individual level, across time, can be explored as outlined in Figure 5.1 (page xcvi) (Curran & Willoughby, 2003; Duncan & Duncan, 2009; Neyer & Lehnart, 2007; Tomarken & Waller, 2005).

5.3 Aims and objectives

The aim of this study is to increase our understanding of the association between PA and balance in an older adult (≥50 years) and to identify a pattern of change in balance performance due to PA. The following objectives were identified:

- To identify indirect measures (functional tests) of balance (cognitive, neuromuscular, and sensory body measures) from the ELSA study that are similar to the measures from the TILDA analysis and that map onto the different body systems required for balance (neuromuscular, cognitive, and sensory) (Table 4.1).
- To carry out CFA (Chapter three, section 3.5) to validate if these indirect measures of balance robustly measure composite measure of balance in community-dwelling older adults (≥50 years) within the ELSA study population across time.

- To gain an understanding of how PA intensity affects balance performance across time using an LGM analysis that incorporates the composite measure of PA previously developed and described in Chapter four which includes inactive, LPA, and MVPA groups; three (indirect) measures of balance with the strongest relationship with balance identified through the CFA analysis; and the covariates of age, sex, activities of daily living (ADL), fall history, pain, alcohol consumption, and education level available from the ELSA study.

5.4 Methods

5.4.1 Study design

The ELSA study design is described in detail in Chapter four (section 4.4.1). In brief, the data used in this study were collected from the ELSA study, a nationally representative study of 12,099 men and women from the English population born on or before February 1952. This analysis uses repeated measures of the same individuals across multiple waves (six waves) over a 10-year period collected from 2004, 2006, 2008, 2010, 2012, 2014 (Steptoe et al., 2013).

5.4.2 Measures

Measures were identified in the raw ELSA data files using a combination of the user guides and derived variables guides for each wave of data collection. These variables were then prepared for analysis in M-plus (version 7.4) in accordance with the M-plus user manual (Muthén & Muthén, 1998-2017) as detailed in (Chapter four, Table 4.1).

5.4.2.1 PA measure

The PA measure is described in Chapter four (section 4.4.2) in detail. In brief, the PA measures is a categorical variable that provides data relating to both PA type and frequency. An LCA approach was used to develop a composite variable of PA intensity consisting of three classes of 'inactive', 'LPA, and 'MVPA' (described in Chapter four, section 4.5.2). These three classes are used in this analysis.

5.4.2.2 Balance

Balance requires a contribution from multiple body systems so seven indirect measures were identified across cognitive, neuromuscular, and sensory systems (as outlined in Chapter four, section 4.4.2) and illustrated in Figure 5.1 (page xcvi). A high score on balance indicates good balance using the scoring system in this study.

5.4.2.3 Covariates

The covariates of age, sex, ADL, falls, pain, alcohol consumption, and education are included in this analysis as these are identified as risk factors for poor balance and are available from the ELSA data. Descriptive statistics for each measure are outlined in Table 5.1 (page lxxxviii).

5.5 Statistical analysis

5.5.1 The measurement model

Firstly, CFA in Mplus (version 7.4; Muthen & Muthen, Los Angeles, CA) was used within a SEM approach to explore whether the indirect measures of balance across different body systems required for balance (Figure 5.1, page xcvi) could be attributed to a composite measure of balance across six waves of data (configural invariance); whether each measure demonstrated equal relationships with the construct of balance across time (metric invariance); and whether differences over time in balance are due to true change in the underlying construct (scalar invariance). Model fit was established using the criteria set out in Chapter three (section 3.5).

5.5.2 The structural model

Having identified those observed measures with the highest factor loading estimates from CFA analysis, a multigroup latent growth model (MG-LGM), was used to understand the trajectory of change in balance over 10 years for inactive, LPA, and MVPA groups (as defined through LCA analysis described in Chapter four, section 4.5) (Figure 5.2, page xcvii). Firstly, to check the variability of responses to the measures, an unconditioned model with no restrictions on the intercept or slope was run. Then, where the variances of the intercept or slope was identified as statistically non-significant ($p>0.05$) restrictions were imposed (the variance was set at zero). Where either the intercept or the slope was statistically significant then the covariates of age, sex, ADL, alcohol consumption, pain, education level, and fall history were included as potential explanatory variables.

A complex analysis option in Mplus was used to account for clustering and stratification. Maximum Likelihood Estimation with robust standard errors (MLR) was used and is robust to non-normality (Enders, 2013; Yaun & Bentler, 2000). Missing data were assumed to be missing at random where systematic differences between the missing and observed values are assumed to be explained by other observed variables (Schafer & Graham, 2002), and MLR utilises a model-based strategy for dealing with missing data which enables all 12,099 participants to be included in analysis.

Model fit was evaluated using a Root Mean Square Error of Approximation (RMSEA) ≤0.05 with an upper limit (90% CI) ≤0.08; a Comparative Fit Index (CFI) ≥0.95; a Tucker Lewis Index (TLI) ≥0.95; and a Standardised Root Mean Square Residual (SRMR) ≤0.08 (Hoyle, 1995). Where the levels of fit indices were not achieved, the modification indices were examined, and where appropriate, adjustments made (Appendix XIII).

5.6 Results

5.6.1 The model of balance

Standardised factor loadings indicated that all observed measures had a statistically significant relationship with the first order latent variables of vision, hearing, steadiness, cognitive function, gait speed, grip strength, and chair rise speed at all six time points as illustrated in Figures 5.3, 5.4, 5.5, 5.6, 5.7, and 5.8. Additionally, all first order latent variables (vision, hearing, cognitive function, steadiness, gait speed, grip strength, and chair rise) had a statistically significant relationship with the second order latent construct or composite measure of balance across all waves. For example, wave two factor loadings showed cognitive function (Est=0.61, SE=0.01), steadiness (Est=0.60, SE=0.02), gait speed (Est=0.78, SE=0.03), grip strength (Est=0.49, SE=0.01), and chair rise speed (Est=1, SE=0.02) had the strongest relationship with the second order latent construct of balance whilst vision (Est=0.41, SE=0.01) and hearing (Est=0.25, SE=0.02) had a weaker relationship. As a result, a residual correlation was introduced between these two variables across all waves because these measures showed a variance not explained by balance.

A series of successive restrictions on the factor loadings for each measure of balance (metric invariance) can be assumed for the factor loadings across all waves, showing that each measure demonstrated an equal relationship with balance over time. A model was then run where the loadings and intercepts were constrained to be equal and scalar invariance could only be assumed for vision and cognitive function across all waves. Therefore, partial invariance can be assumed, and as factor loadings were constrained successfully across all waves, and two intercepts constrained across all waves, then valid inferences about the differences between latent factor means can be made (Byrne, Shavelson &

Muthen, 1989) (Table 5.2, page lxxxix).

Standardised results showed that balance was highly correlated across all six waves of data (0.96, 1, 0.99, 0.99, and 0.99 respectively). The mean of balance increased slightly between wave one and two (2.15 and 2.17 respectively), but then declined over subsequent years (1.88, 1.84, 1.60, 1.52).

CFA analyses (outlined above) identified that the individual observed measures of gait speed, cognitive function, and steadiness had the strongest relationship with balance and so single observed measures of each were included in an LGM. Firstly, an unconditioned model for each measure of gait, cognitive function, and steadiness was run (e.g. no restrictions on intercepts or slopes) with the covariates of age, sex, history of falls, ADL, pain, alcohol consumption level, and education level (Table 5.3, page xc). Where non-significance was identified on the intercept or slope then restrictions were imposed. In a number of models, the slopes were fixed to zero as the lack of inter-individual variability in the slopes had results in a nonconvergent solution. The fit statistics for each measure (gait speed, cognitive function, and steadiness) are outlined in Table 5.4 (page xcii). In general, all models provided a satisfactory description of the data where chi-squared tests of model fit were not statistically significant (0.05 level), the CLI and TFI measures were above 0.95, the RMSEA was below 0.05 with an upper bound below 0.08. The following provides the detail for each measure.

5.6.2 The LGM for gait speed

The model for the measure of gait speed consisted of five timepoints (waves two, three, four, five, and six), and was fitted to the data by setting the slope value for each time point to values of 0, 1, 2, 3, and 4 respectively (linear model). The slope for each PA group was allowed to be freely estimated. Where the inter-individual differences were not statistically significant at the 0.05 level then the effect of the covariates/explanatory variables was not explored.

The results showed that there was statistically significant variability in the speed achieved on the measure of gait between PA groups where the MVPA group (intercept Est=1.12, SE=0,05) scored higher than the LPA group (intercept Est=0.96; SE=0.12), and where the LPA group scored higher than the inactive group (intercept Est=0.84, SE=0.08). This inter-individual variability was explored in terms of the effects of all covariates, but only age (Est=-0.14; SE=0.04) was statistically significant and therefore contributed to the change in the measure of gait.

The extent of inter-individual change in the measure of gait speed is small, but statistically significant in inactive (slope Est=-0.09, SE=0.04), LPA (slope Est= -0.07, SE=0.02), and MVPA (slope Est=-0.09, SE=0.02) groups. Therefore, the change in the measure of gait over time varies across individuals within the PA groups, and to understand if the effects were the same within different age categories, three age categories of 'young-old' (60-70 years), 'old-old' (70-80 years), and 'old-older' (80-90 years) were used (Figure 5.13, page cviii). The graph shows that the MVPA group achieved faster gait speed in ≤70 years, but the LPA group achieved better gait speed ≥70 years. It may be possible that this result is due to a blunted response in the LPA group, or perhaps the LPA group are achieving much more volume on LPA and so the result may be due to volume rather than intensity of PA. The MVPA group showed a slower rate of decline than other groups, and this is further supported using the calculated change in the predicted value (Table 5.6, page xciv) which shows that the change in gait speed was small across all PA groups, and the rate of change in the MVPA group is positive compared with LPA and inactive groups indicating a slower rate of decline (Appendix XV).

The correlation between the intercept and slope were significant for inactive (Est=-0.01, SE=0.01) and MVPA (Est=-0.02, SE=0.003) groups indicating that those participants who performed better on the gait test tended to show less decline in gait speed over time within groups. The correlation for LPA (Est=-0.01, SE=0.01) was not significant suggesting that those participants who performed better on gait tended to show more decline in gait speed over time within groups.

5.6.3 The LGM for steadiness

The self-reported measure of steadiness consisted of three timepoints (wave one, two and four) and the model was fitted to the data by setting the slope value for the first and second timepoints to values of 0, and 1 respectively. The remaining slope was allowed to be freely estimated. The slope across all PA groups was fixed to zero as lack of inter-individual variability on the slope had resulted in a nonconvergent solution. Where inter-individual differences were not statistically significant at the 0.05 level then the effect of the covariates was not explored.

The results showed that there is statistically significant variability between PA groups when responding to the question on steadiness where the MVPA group (Est=3.91, SE=0.18) scored themselves higher than the LPA (Est=3.33, SE=0.50), and the LPA group scored themselves higher than the inactive group (Est=2.17, SE=0.86). The inter-individual variability in terms of the effects of covariates was explored and only fall history and ADL were statistically significant across all groups on the initial scores (intercept) and therefore contributed to the change in steadiness. Additionally, pain was significant for LPA and MPVA groups. For example, a history of falls in the last 2 years, difficulty in ADL, and the presence of pain, resulted in lower scores on steady and therefore an indication of poor balance (Figure 5.10-5.13, pages cv-cviii).

The extent of inter-individual change in the measure of steadiness is small and not statistically significant at the 0.05 level in either inactive (slope Est=-0.2, SE=0.46), LPA (slope Est=-0.03, SE=0.30), or MVPA (slope Est=-0.07, SE=0.15). Therefore, the change occurring in steadiness over time is similar across individuals within each PA group. As a result, the effect of the covariates was not explored, but to understand if the effects were the same in all age categories three age categories 'young-old' (60-70 years), 'old-old' (70-80 years), and 'old-older' (80-90 years) were explored (Figure 5.13, page cviii). The graph shows that the MVPA group rated their steadiness higher than any other PA group within each age category, and the MVPA group showed a slower rate of decline than other groups. This is further supported using the calculated change in the predicted value (Table 5.6, page xciv) which shows that the change in steadiness was small across all PA groups, and the rate of change in the MVPA group is positive compared with LPA and inactive groups indicating a slower rate of decline (Appendix XVI). The correlation between the intercept and slope was not estimated as the slope was fixed to zero for convergence.

5.6.4 The LGM for cognitive function (COG)

The measure of cognitive function consisted of four timepoints (wave one, two, four, and five), and the model for cognitive function was fitted to the data by setting the slope value to each timepoint to 0, 1, 2, and 3 respectively. The slope for the inactive group was set to zero as lack of inter-individual variability on the slope had resulted in a nonconvergent solution. Where inter-individual differences were not statistically significant at the 0.05 level then the effect of the covariates was not explored.

The results showed that there is statistically significant variability between PA groups when responding to the question on cognitive function where the LPA group (Est=23.59, SE=2.06) scored better than either MVPA group (Est=22.33, SE=1.84) or inactive (Est=15.13, SE=3.50) groups. This inter-individual variability was using the covariates. Of these measures pain (Est=2.12; SE=0.96) was statistically significant on the intercept for the inactive group where the absence of pain was associated with improved score on cognitive function; age (Est=-0.61; SE=0.53), ADL (Est=1.91; SE=0.82), fall history (Est=-1.58; SE=0.81) and education level (Est=-0.99; SE=0.18) were statistically significant for the LPA group, where increased age, ADL disability, a history of falls, and lower education level is associated with lower cognitive function; and age (Est=-1.72; SE=0.43), and education level (Est=-0.52; SE=0.12) were statistically significant for MVPA where older adults with lower education level is associated with lower scores on cognitive function (Figure 5.14, page cix).

The extent of inter-individual change in the measure of cognitive function was not statistically significant at the 0.05 level in either inactive (slope Est=1.02; SE=1.23), LPA (slope Est=-1.33; SE=1.06), or MVPA (slope Est=0.91; SE=0.73). Therefore, the change occurring in cognitive function over time is similar across individuals within each PA group. As a result, the effect of the covariates was not explored, but to understand if the effects were the same in all age categories three age categories ‘young-old’ (60-70 years), ‘old-old’ (70-80 years), and ‘old-older’ (80-90 years) were explored (Figure 5.13, page cviii). The graph shows that ≤70 years the MVPA group responded better to the measure of cognitive function than other PA groups, but ≥70 years the LPA group responded better than other PA groups and improved their score over time compared with other PA groups that declined. Also, the rate of decline was slower for the MVPA group than the inactive group (Table 5.7, page xcv; Appendix XVII).

The correlation between the intercept and slope was positive in the LPA group (Est=1.18; SE=0.99) indicating that if scores are low on the intercept then change is more rapid. In contrast, the correlation was negative in the MVPA group

(Est=-0.13, SE=0.16) indicating that those who performed better on cognitive function tended to show less change over time.

5.7 Discussion

5.7.1 Principal findings

CFA analysis supports that multiple indirect functional measures of cognitive function, of the neuromuscular system (chair rise test, gait speed test, handgrip strength, and steadiness), and of the sensory system (self-rated vision and hearing) collectively provide a model for balance assessment in an older population in England (Horak, 1995; Sibley et al., 2015). The measures of gait speed, self-rated steadiness, and cognitive function were found to have the strongest relationship with balance.

An LGM analysis of the effects of PA intensity (inactive, LPA, and MVPA) on the measures of gait speed, self-rated steadiness, and cognitive function suggests a dose-response relationship between PA and balance (Powell et al., 2011). For example, the MVPA group had better scores than LPA and inactive groups for measures of gait speed and steadiness. The LPA group had better scores on all three measures than the inactive group (controlling for age). Furthermore, MVPA was found to slow the rate of decline in all measures, excluding cognitive function where LPA maintained balance. These findings support existing clinical evidence (Bandeen-Roche et al., 2015; Brigola et al., 2015; Bucknix et al, 2015; Cadore et al., 2013; Chen et al., 2014; 2015; Gillespie et al., 2012; Howe et al., 2011; Karlsson et al., 2013; Stubbs et al., 2015; Theou et al., 2011; WHO, 2007), and existing healthcare policy and promotion (CSP), 2018; DoH, 2011; NICE, 2013; PHE, 2017, 2018; RCP, 2015; WHO, 2007, 2010, 2015, 2018) that suggests that higher intensity activity such as MVPA benefits balance.

Additionally, the findings extend our understanding by providing longitudinal analysis of the benefits of LPA on balance performance (Bauman et al., 2016; NICE, 2013; PAGAC, 2018; Schmidt et al., 2017). The analysis of the effect of age showed that older adults ≤ 70 years in the MVPA group had faster gait speed, and higher scores on cognitive function and self-rated steadiness. However, older adults ≥70 years in the LPA group had better gait speed and cognitive function than the MVPA group (Figures 5.9, 5.14). This suggests that older adults (≥70 years) benefit more from LPA for balance performance perhaps because of a ceiling effect where even small increases in PA can benefit this older age group (Powell et al., 2011). Additionally, the analysis identified that covariates such as pain, ADL disability, and fall history were statistically significant for older adults ≥70 years in LPA and inactive groups indicating that these groups were perhaps physically incapable of engaging in MVPA (Ainsworth et al., 2011; Arnardottir et al., 2013; Franco et al., 2015) and so had more potential for improvement from any increase in activity.

In summary, balance declines across time and older adults ≤70 years may benefit more from MVPA, but older adults ≥70 years may experience more benefit from LPA as they may be able to sustain this activity due to their physical capability, and therefore may also have more potential for improvement in measures of balance. Therefore, policy should encourage LPA activity for higher risk of falling adults (≥70 years) and encourage MVPA for adults at less risk of falling (≤70 years).

5.7.2 Methodological quality

Chapter two highlighted that research exploring the effects of PA and balance were, in the main, of low to moderate quality, and this study addresses some of the methodological issues highlighted as highlighted by Chapter three. For example, the ecological validity is high as this study uses a large representative sample of data over a ten-year period from a robustly designed longitudinal study of ageing (ELSA); a SEM statistical approach which minimises measurement error and validates the measures for older adults living in England; and the analysis uses MLR, thus allowing all participant data to be used, therefore reducing the bias of the results (Enders, 2013).

5.7.3 Strengths and weaknesses

A key strength of this study is that it identified those functional tests with the strongest relationship with balance through MG-CFA, but to fully understand change in balance over time, then balance as a multi-system approach should be considered rather than individual measures which improves the robustness of the findings by addressing measurement error (Duncan & Duncan, 2009; Sibley et al., 2015). Furthermore, whilst LGM analysis included objective measures of gait speed and cognitive function, it also included the self-rated measure of steadiness. As highlighted in Chapter one, self-reported measures are subject to bias (Dyrstad et al., 2014; Murphy, 2009; Saelens et al., 2012), and therefore, the findings from the analysis steadiness should be considered with caution.

Furthermore, this study uses LGM to analyse longitudinal data collected over a 10-year period at six different time points to understand how different intensities of PA affect balance in older adults over time (Duncan & Duncan, 2009). The analysis consisted of a two-stage approach where the composite measure of PA was developed using LCA (Chapter four), and then the different PA groups were included in the LGM analysis. Future research should integrate the LCA into the overall LGM analysis to improve the robustness of the analysis as the integrated approach ensures that the latent classes are relevant to the whole sample.

5.7.4 Future research and clinical implications

This study addresses research recommendations and clinical implications presented in Chapter three where the results can guide the development of more robust balance screening measures (Horak, 1995; Sibley et al., 2015) to assess risk of falls more accurately, and target interventions more appropriately.

This study uses the observed measures of balance to explore individual differences, but future studies should explore balance as a multi-dimensional construct in the context of LCA to improve the validity and reliability of the model. Additionally, whilst this study identifies the benefits of LPA for older adults ≥70 years, future research should explore the amount (minimum and maximum dosage) of LPA required to realise the benefits for balance so that future guidelines and policy can provide more explicit guidance.

5.8 Conclusion

The neuromuscular measures of chair rise test, gait speed test, handgrip strength, and steadiness, the sensory measures of vision and hearing, and the cognitive measures identified from the ELSA study collectively provide a model for balance assessment in an older population in England. LGM analysis identified that higher intensity PA is better than lower intensity PA for older adults ≤70 years and can also reduce the decline in balance measures over time. Furthermore, LPA has more benefits for older adults aged ≥70 years for balance measures.

Table 5.1
Descriptive statistics of observed measures of balance for population and by latent class

Observed measure	Total population	Contribution by latent class: Inactive	Low intensity	Mod/vig intensity
Gripa	2776	390	673	1713
Gripb	2773	390	672	1711
Gripc	2771	389	671	1710
Chair	1636	230	415	991
Gaita2c	2383	244	570	1569
Gaitb2c	2366	242	560	1564
Steady	2985	432	719	1834
Steady2	2984	431	719	1834
Eyea	2986	432	719	1835
Eyeb	2981	431	717	1833
Eyec	2981	431	717	1833
Eara	2986	432	719	1835
Earb	2986	432	719	1835
Cfra	2984	432	718	1834
Cfrb	2982	430	718	1834
Cfmd	2953	415	714	1824
Cfre	2958	420	714	1824
Cffc	2982	431	717	1834

Table 5.2
Fit statistics for the latent construct of balance waves one to six

Models	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	Value	df	P-value	Estimate	90 % C.I.	CFI	TLI	Value
1.3 Latent construct of balance waves 1-6 (configural variance)	998102.48	1000930.92	8756.06	3011	0.0000	0.01	(0.01, 0.01)	0.98	0.97	0.05
1.4 Latent construct of balance waves 1-6 (metric invariance)	998232.46	1000778.78	8787.34	3050	0.0000	0.01	(0.01, 0.01)	0.98	0.97	0.05
1.5 Latent construct of balance waves 1-6 (scalar invariance)	999543.83	1001699.53	9998.40	3104	0.0000	0.02	(0.01, 0.02)	0.97	0.97	0.05

Table 5.3
Latent growth model fit statistics for each observed measure of balance

Observed measure of balance	Information Criteria		Chi-square			Contribution of Chi-square PA class			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	Value	df	P-value	Inactive	LPA	MVPA	Estimate	90% C.I.	CFI	TLI	Value
Gripa	3253.28	3466.72	13.26	9	0.1510	1.56	6.74	4.96	0.02	(0.00, 0.05)	0.99	0.99	0.01
Gripb	2919.63	3133.02	4.56	9	0.8711	0.97	1.76	1.82	0.00	(0.00, 0.02)	1.00	1.00	0.00
Gripc	2876.47	3089.82	9.68	9	0.3773	1.46	0.71	7.50	0.01	(0.00, 0.04)	0,99	1.00	0.00
Chair	14117.73	14312.14	9.04	9	0.4332	2.19	5.15	1.69	0.00	(0.00, 0.05)	1.00	1.00	0.02
Gaita2c	7574.94	7817.53	147.45	48	0.0000	29.99	37.59	79.86	0.05	(0.04, 0.06)	0.93	0.92	0.04
Gaitb2c	6714.75	6957.05	161.73	48	0.0000	26.97	49.26	85.51	0.06	(0.05, 0.06)	0.93	0.91	0.05
Steady2	9031.32	9247.56	14.01	9	0.1218	9.105	1.97	2.94	0.02	(0.00, 0.05)	0.99	0.99	0.01
Steady	16234.34	16450.39	31.29	9	0.0003	5.88	14.17	11.24	0.05	(0.03, 0.07)	0.99	0.96	0.02
Eyea	42556.74	42826.82	154.38	72	0.0000	62.53	23.92	68.21	0.03	(0.03, 0.04)	0.98	0.98	0.03
Eyeb	40867.69	41137.69	163.68	72	0.0000	39.38	47.92	76.37	0.04	(0.03, 0.04)	0.98	0.97	0.03
Eyec	41777.18	42047.18	158.68	72	0.0000	49.92	46.59	62.18	0.04	(0.03, 0.04)	0.98	0.97	0.04
Eara	42167.60	42437.67	102.09	72	0.0113	29.77	21.94	50.39	0.02	(0.01, 0.03)	0.99	0.99	0.02
Earb	16858.03	17128.10	186.60	72	0.0000	67.98	35.73	82.88	0.04	(0.03, 0.05)	0.98	0.98	0.04

Table 5.3 (continued)
Latent growth model fit statistics for each observed measure of balance

Observed measure of balance	Information Criteria		Chi-square			Chi-square PA class			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	Value	df	P-value	Inactive	LPA	MVPA	Estimate	90% C.I.	CFI	TLI	Value
Cfra	17647.27	17899.32	50.53	48	0.3738	11.82	16.91	21.80	0.01	(0.00, 0.02)	0.99	0.99	0.04
Cfrb	49432.63	49684.64	80.57	48	0.0022	10.30	23.28	46.99	0.03	(0.02, 0.04)	0.99	0.99	0.04
Cfmd	35491.28	35724.92	50.21	27	0.0043	12.42	13.19	24.60	0.03	(0.02, 0.04)	0.96	0.94	0.03
Cfre	48359.10	48610.77	65.89	48	0.0441	24.22	17.61	24.06	0.02	(0.00, 0.03)	0.99	0.99	0.02
Cffc	68762.26	68996.27	49.71	27	0.0049	17.47	14.17	18.07	0.03	(0.02, 0.04)	0.99	0.99	0.03

Table 5.4
Fit statistics for observed measures of gait, cognitive function, and steadiness

Models	**Information Criteria**		**Chi squared**			**RMSEA**		**CFI/TLI**		**SRMR**
	Akaike (AIC)	Bayesian (BIC)	Value	df	P-value	Estimate	90 % C.I.	CFI	TLI	Value
LGM for observed measure of Steady	4755.62	5026.45	30.99	31	0.4669	0.00	(0.00, 0.05)	1.00	1.00	0.04
LGM for observed measure of Cognitive function (cffc)	16968.15	17234.47	87.36	68	0.0569	0.03	(0.00, 0.05)	0.98	0.97	0.04
LGM for observed measure of Gait	1660.45	1963.28	113.36	95	0.0964	0.03	(0.00, 0.05)	0.96	0.94	0.07

Table 5.5
Predicted values and rate of change in predicted values for gait speed across inactive, LPA, and MVPA groups

Wave	Inactive group	LPA group	MVPA group
Wave 1: predicted value	0.96+-0.10 (0) =0.96	0.84+ -0.07 (0) = 0.84	1.12+-0.09 (0) = 1.12
Change	0	0	0
Wave 2: predicted value	0.96+-0.10 (1) = 0.86	0.84+ -0.07 (1) = 0.77	1.12+-0.09 (1) = 1.03
Change	0.96 - 0.86= 0.10	0.84 - 0.77 = 0.07	1.12 - 1.03 = 0.09
Wave 3: predicted value	0.96+-0.10 (2) =0.76	0.84+ -0.07 (2) = 0.70	1.12+-0.09 (2) =0.94
Change	0.96 - 0.76= 0.20	0.84 - 0.70 = 0.14	1.12 - 0.94= 0.18
Wave 4: predicted value	0.96+-0.10 (3) =0.66	0.84+ -0.07 (3) = 0.63	1.12+-0.09 (3) = 0.85
Change	0.96 - 0.66= 0.30	0.84 – 0.63 = 0.21	1.12 - 0.85 = 0.27
Wave 5: predicted value	0.96+-0.10 (4) = 0.56	0.84+ -0.07 (4) = 0.56	1.12+-0.09 (4) = 0.76
Change	0.96 - 0.56 = 0.40	0.84 - 0.56 = 0.28	1.12 - 0.76 = 0.36

Note: Predicted value of slope = (intercept + slope) x (timepoint estimate value); Change in predicted value = (intercept – predicted value).

Table 5.6
Predicted values and rate of change in predicted values for steadiness across inactive, LPA, and MVPA groups

Wave	Inactive group	LPA group	MVPA group
Wave 1: predicted value	2.23+-0.02 (0) = 2.23	3.43+-0.03 (0) = 3.43	3.91+-0.08 (0) = 3.91
Change	0	0	0
Wave 2: predicted value	2.23+-0.02 (1) = 2.21	3.43+-0.03 (1) = 3.40	3.91+-0.08 (1) = 3.83
Change	2.23-2.21= 0.02	3.43-3.40= 0.03	3.91-3.83= 0.08
Wave 4: predicted value	2.23+-0.02 (2) = 2.19	3.43+-0.03 (2) = 3.37	3.91+-0.08 (2) = 3.75
Change	2.23-2.19= 0.04	3.43-3.37=0.06	3.91-3.75= 0.16

Note: Predicted value of slope = (intercept + slope) x (timepoint estimate value)
Change in predicted value = (intercept – predicted value).

Table 5.7
Predicted values and rate of change in predicted values for COG across inactive, LPA, and MVPA groups

Wave	Inactive group	LPA group	MVPA group
Wave 1: predicted value	15.13+1.02 (0) = 15.13	23.59+ -1.33 (0) = 23.59	22.33+0.91 (0) = 22.33
Change	0	0	0
Wave 2: predicted value	15.13+1.02 (1) = 16.15	23.59+ -1.33 (1) = 22.26	22.33+0.91 (1) = 23.24
Change	15.13 - 16.15 = -1.02	23.59 – 22.26 = 1.33	22.33 - 23.24 = -0.91
Wave 4: predicted value	15.13+1.02 (2) = 17.17	23.59+ -1.33 (2) = 20.93	22.33+0.91 (2) = 24.15
Change	15.13 - 17.17 = -2.04	23.59 – 20.93 = 2.66	22.33 – 24.15 = -1.82
Wave 5: predicted value	15.13+1.02 (3) = 18.19	23.59+ -1.33 (3) = 19.60	22.33+0.91 (3) = 25.06
Change	15.13 – 18.19 = -3.06	23.59 - 19.60 = 3.99	22.33 – 25.06 = -2.73

Note: Predicted value of slope = (intercept + slope) x (timepoint estimate value)
Change in predicted value = (intercept – predicted value).

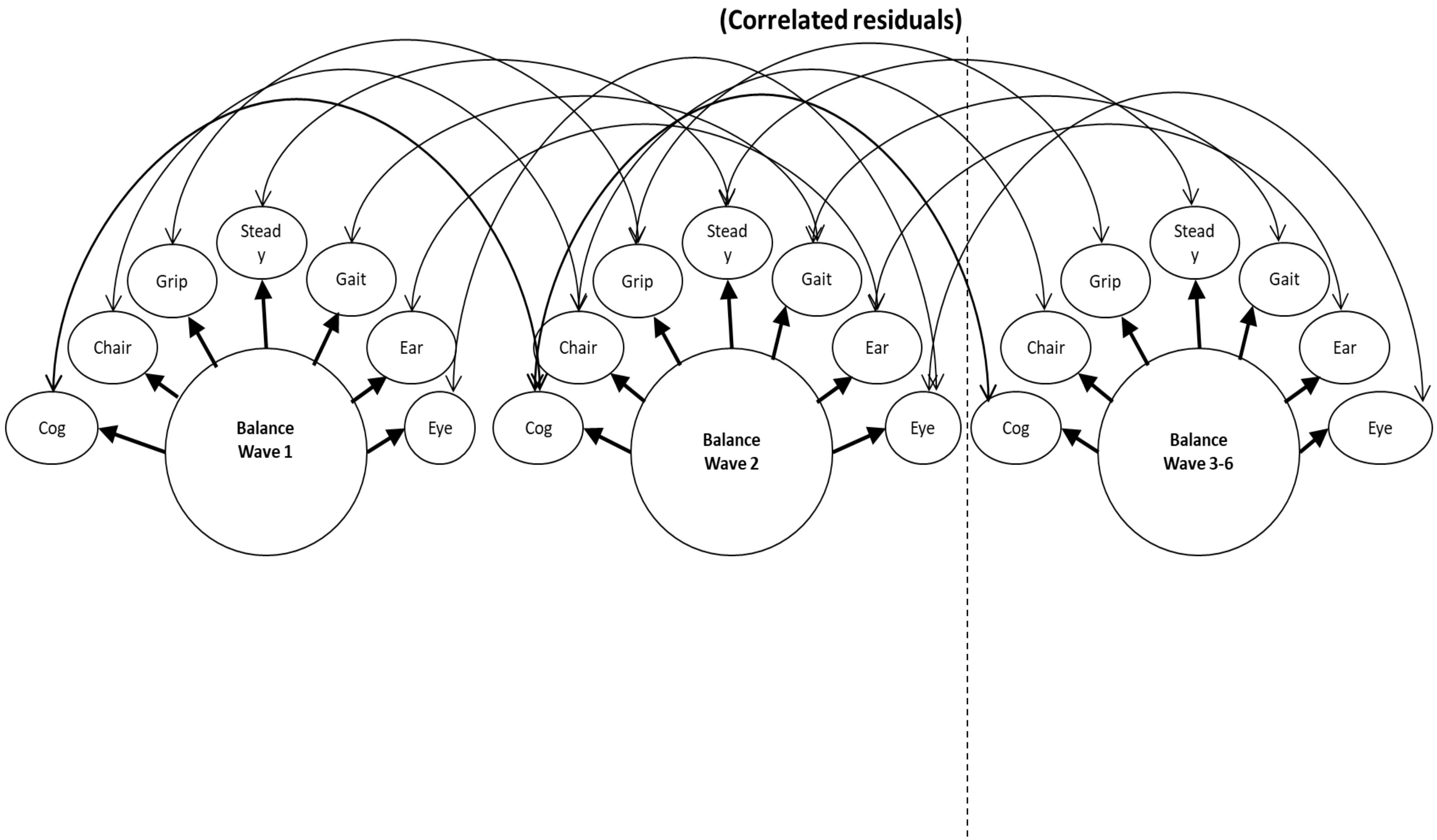


Figure 5.1
Model of balance using indirect measures (neuromuscular, cognitive, and sensory systems) across six waves from the ELSA study

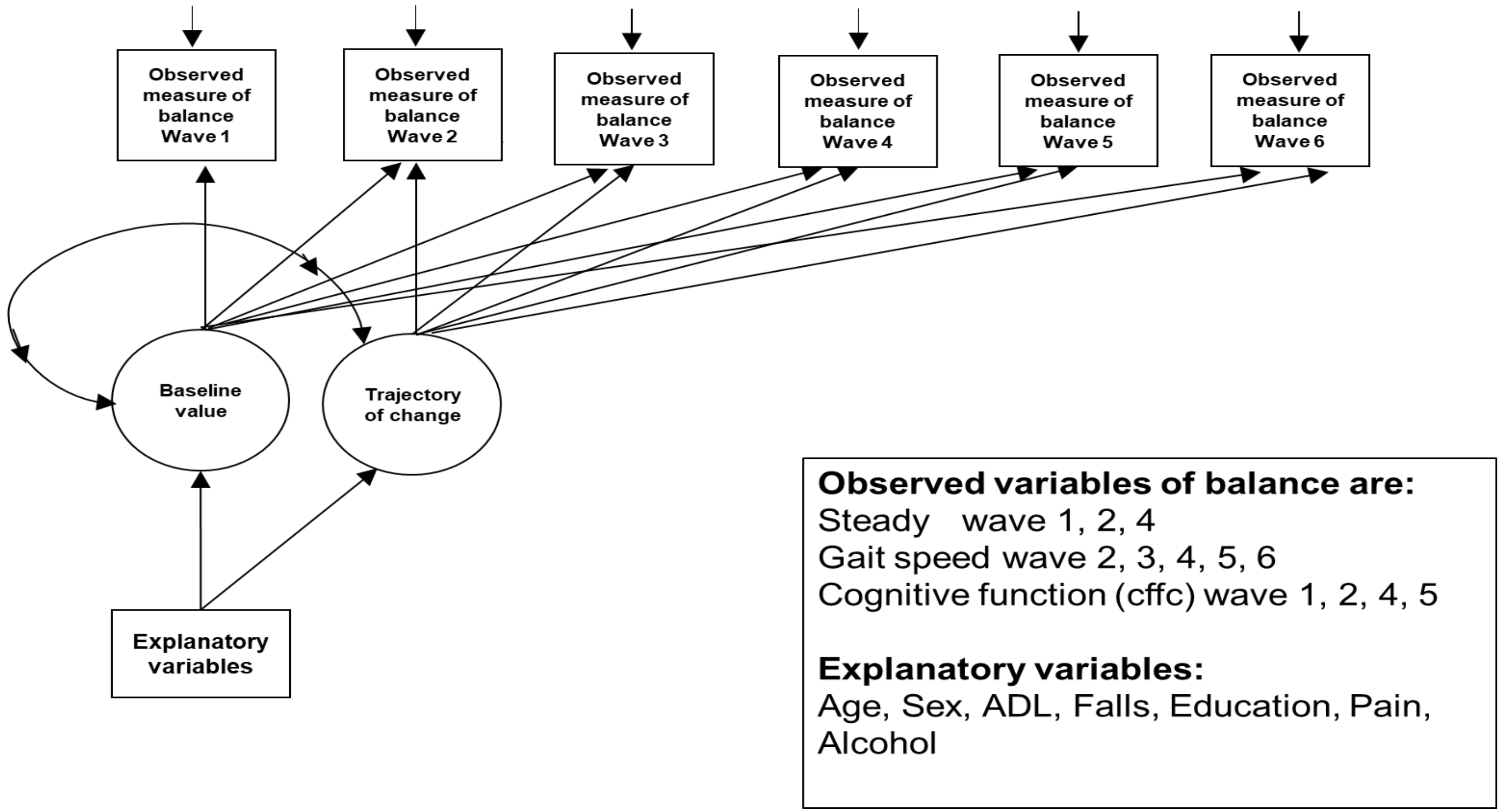
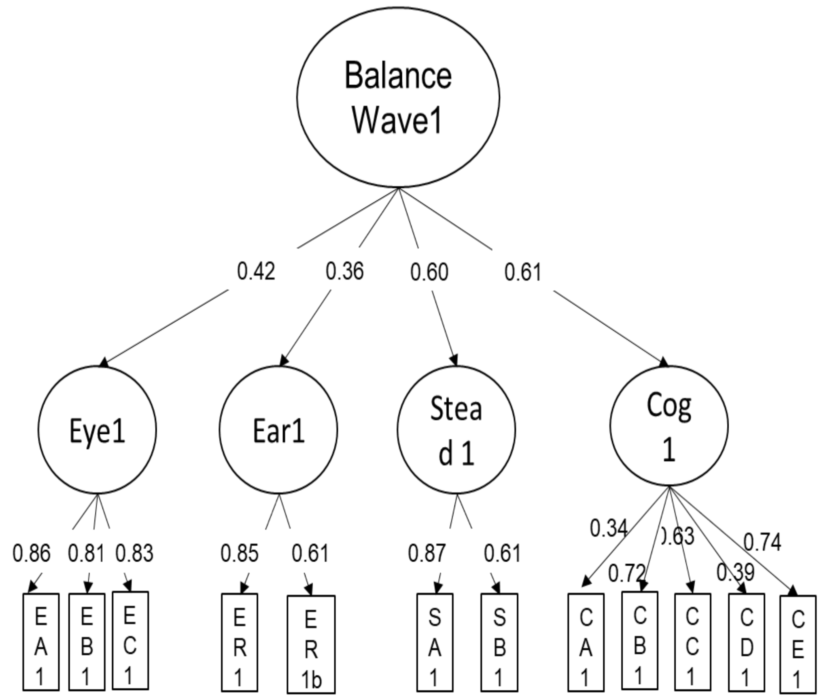


Figure 5.2
Latent growth model using measure of Physical Activity (PA) (Inactive, Low, and Moderate/vigorous intensity PA) and observed balance measures

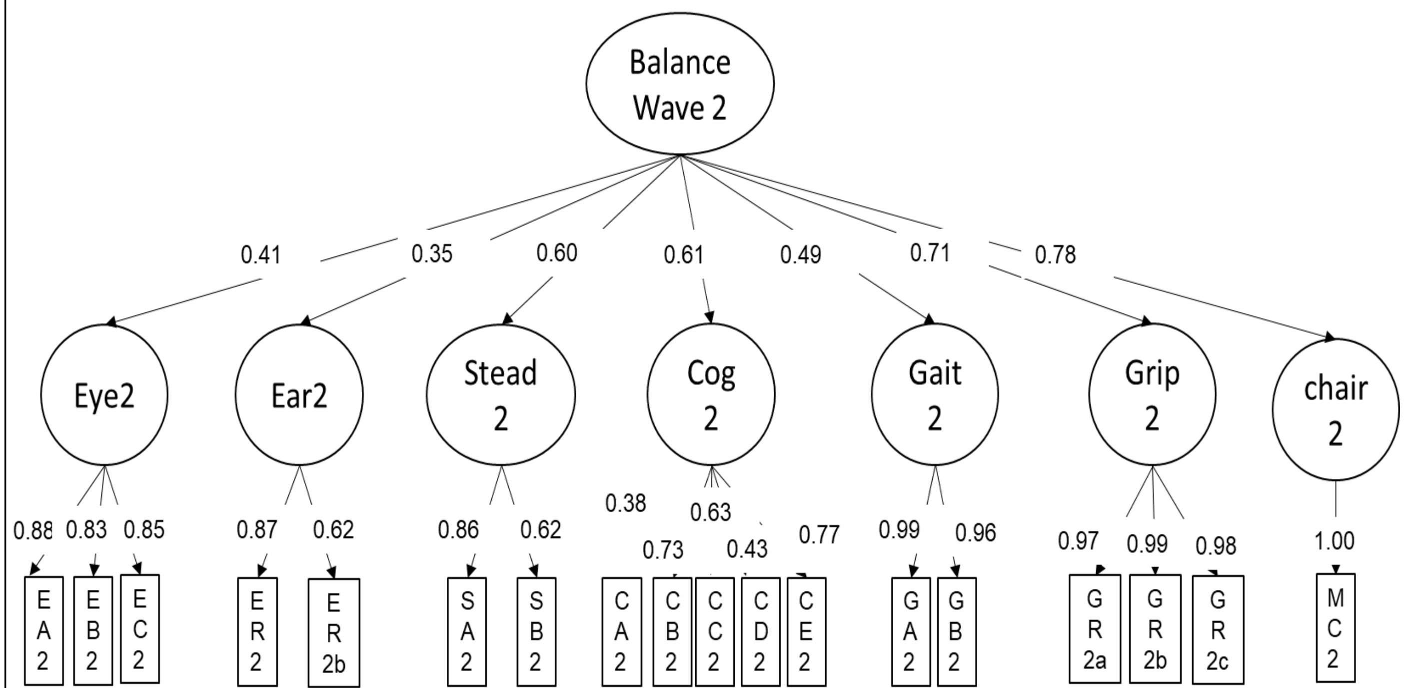
Key:
(x = wave)
EAx = Eyeax
EBx = Eyebx
ECx = Eyecx
Erx = Earax
ERxb=Earbx
Sax = Steadyax
Sbx = Steadybx
Cax = cfrax
Cbx = cfrbx
Ccx = cfrcx
Cdx = cfmdx
Cex = cfrex
Gax = gaitax
Gbx = gaitbx
Grxa = gripax
Grxb = gripbx
Grxc = gripcx



Note: residual correlations and errors are excluded from the diagram

Figure 5.3
Results for first factor loadings for balance at wave 1

Key:
(x = wave)
EAx = Eyeax
EBx = Eyebx
ECx = Eyecx
Erx = Earax
ERxb=Earbx
Sax = Steadyax
Sbx = Steadybx
Cax = cfrax
Cbx = cfrbx
Ccx = cfrcx
Cdx = cfmdx
Cex = cfrex
Gax = gaitax
Gbx = gaitbx
Grxa = gripax
Grxb = gripbx
Grxc = gripcx

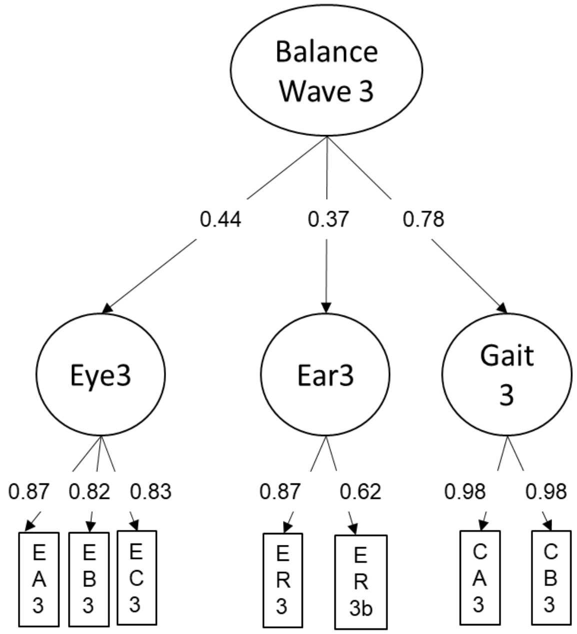


Note: residual correlations are excluded from the diagram

Figure 5.4
Standardised first factor loadings for balance at wave 2

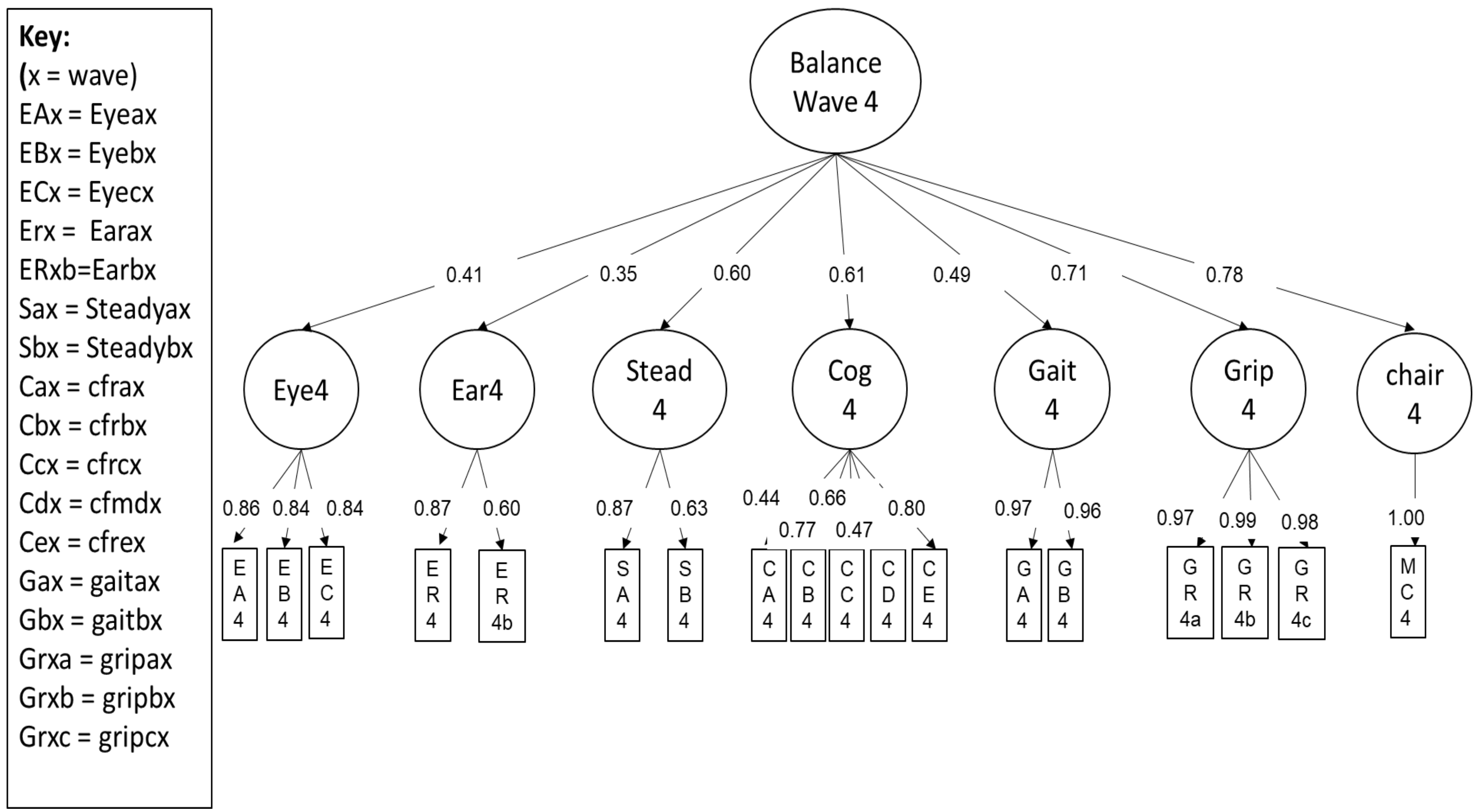
c

Key:
(x = wave)
EAx = Eyeax
EBx = Eyebx
ECx = Eyecx
Erx = Earax
ERxb=Earbx
Sax = Steadyax
Sbx = Steadybx
Cax = cfrax
Cbx = cfrbx
Ccx = cfrcx
Cdx = cfmdx
Cex = cfrex
Gax = gaitax
Gbx = gaitbx
Grxa = gripax
Grxb = gripbx
Grxc = gripcx



Note: residual correlations and errors are excluded from the diagram

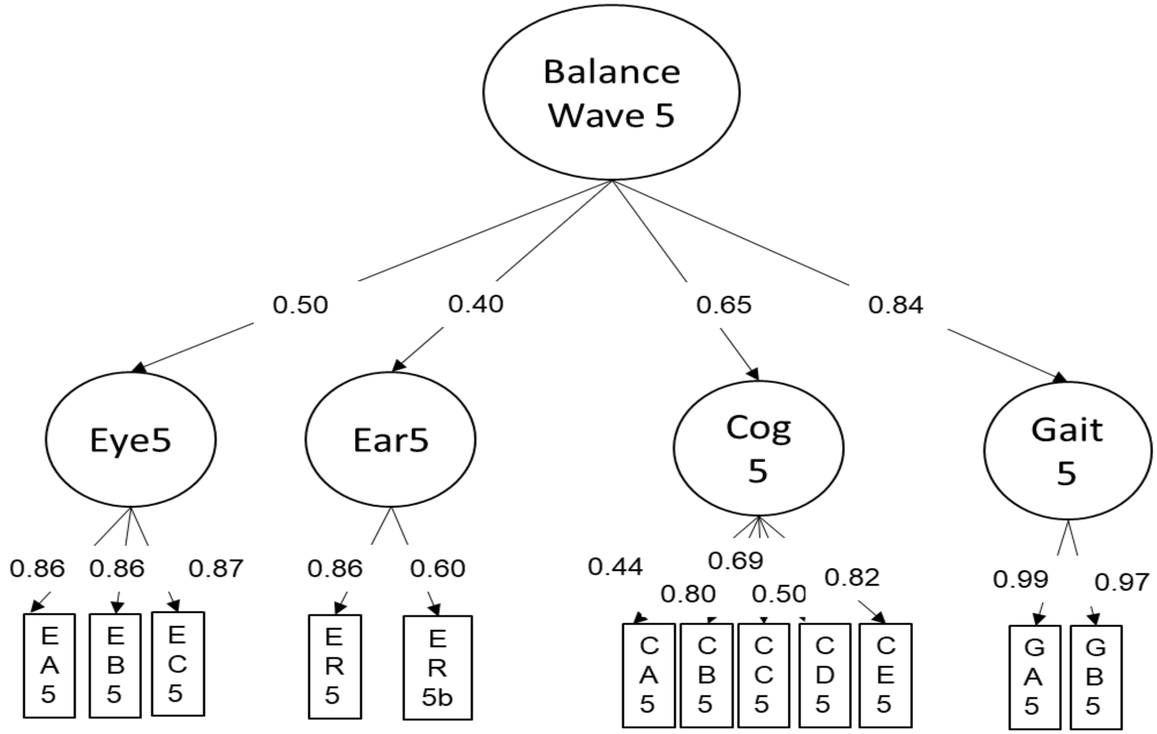
Figure 5.5
Standardised first factor loadings for balance at wave 3



Note: residual correlations are excluded from the diagram

Figure 5.6
Standardised first factor loadings for balance at wave 4

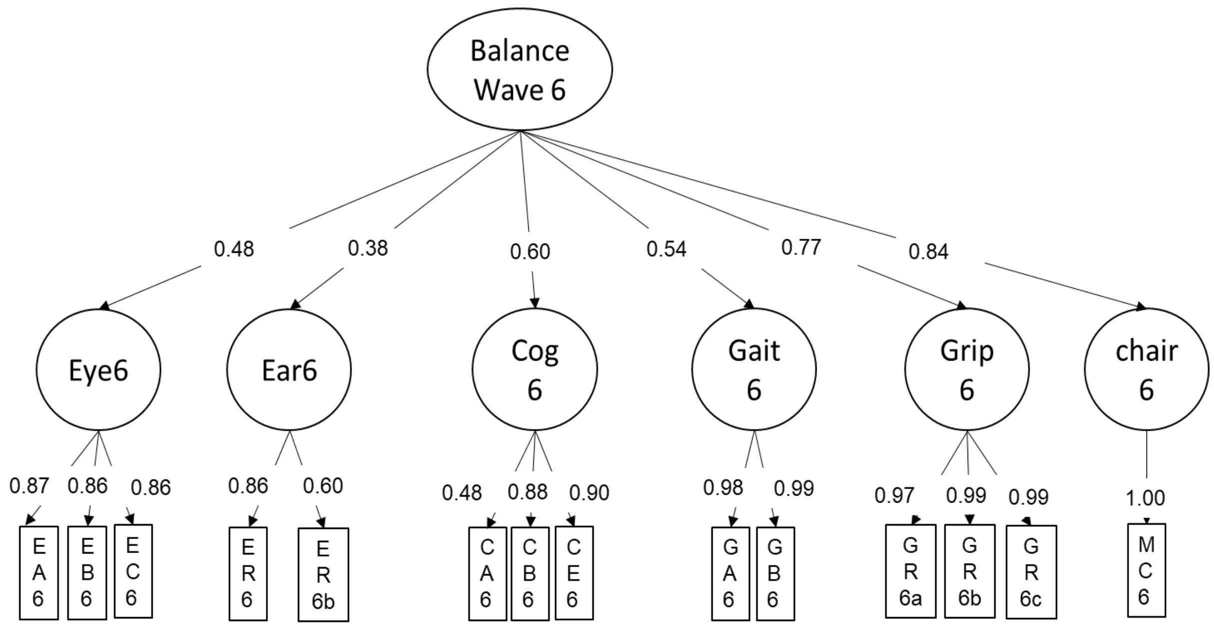
Key:
(x = wave)
EAx = Eyeax
EBx = Eyebx
ECx = Eyecx
Erx = Earax
ERxb=Earbx
Sax = Steadyax
Sbx = Steadybx
Cax = cfrax
Cbx = cfrbx
Ccx = cfrcx
Cdx = cfmdx
Cex = cfrex
Gax = gaitax
Gbx = gaitbx
Grxa = gripax
Grxb = gripbx
Grxc = gripcx



Note: residual correlations are excluded from the diagram

Figure 5.7
Standardised first factor loadings for balance at wave 5

Key:
(x = wave)
EAx = Eyeax
EBx = Eyebx
ECx = Eyecx
Erx = Earax
ERxb=Earbx
Sax = Steadyax
Sbx = Steadybx
Cax = cfrax
Cbx = cfrbx
Ccx = cfrcx
Cdx = cfmdx
Cex = cfrex
Gax = gaitax
Gbx = gaitbx
Grxa = gripax
Grxb = gripbx
Grxc = gripcx



Note: residual correlations and errors are excluded from the diagram

Figure 5.8
Standardised first factor loadings for balance at wave 6

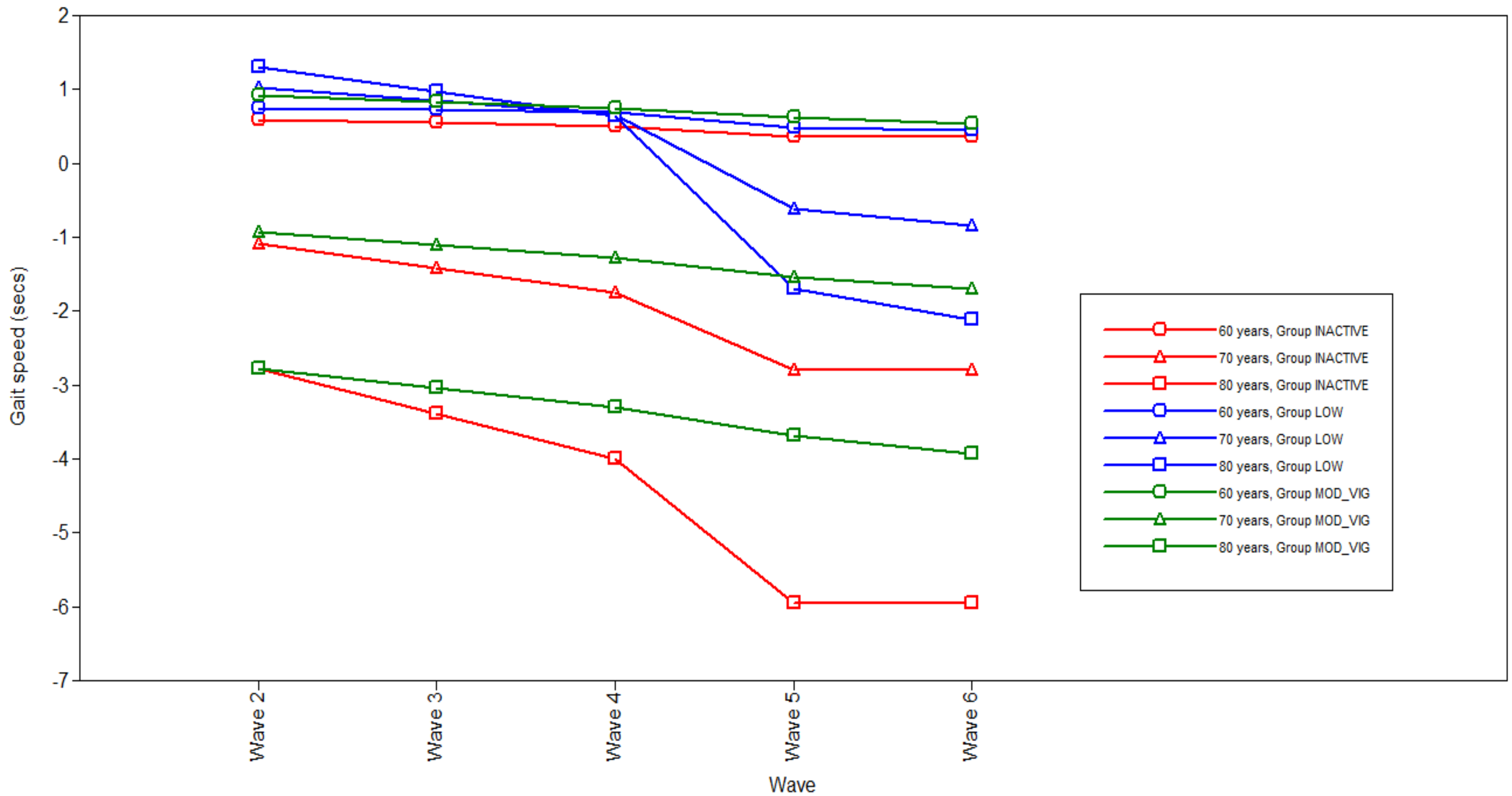


Figure 5.9
The trajectory of change in inactive, low, and moderate/vigorous PA groups on the measure of gait speed controlling for all variables by age

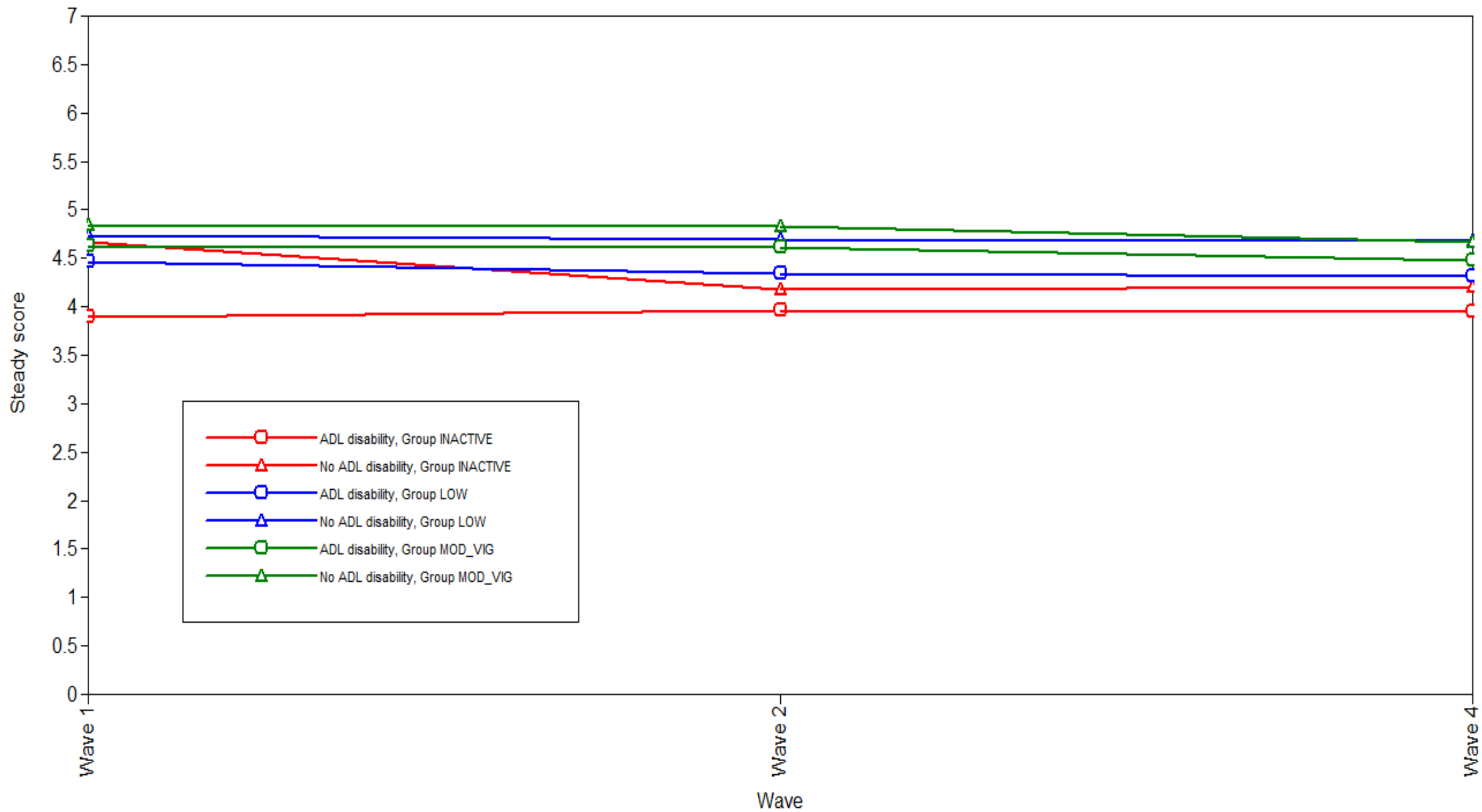


Figure 5.10
The effect of Activities of Daily Living (ADL) on the measure of steadiness for inactive, low, and moderate/vigorous PA groups

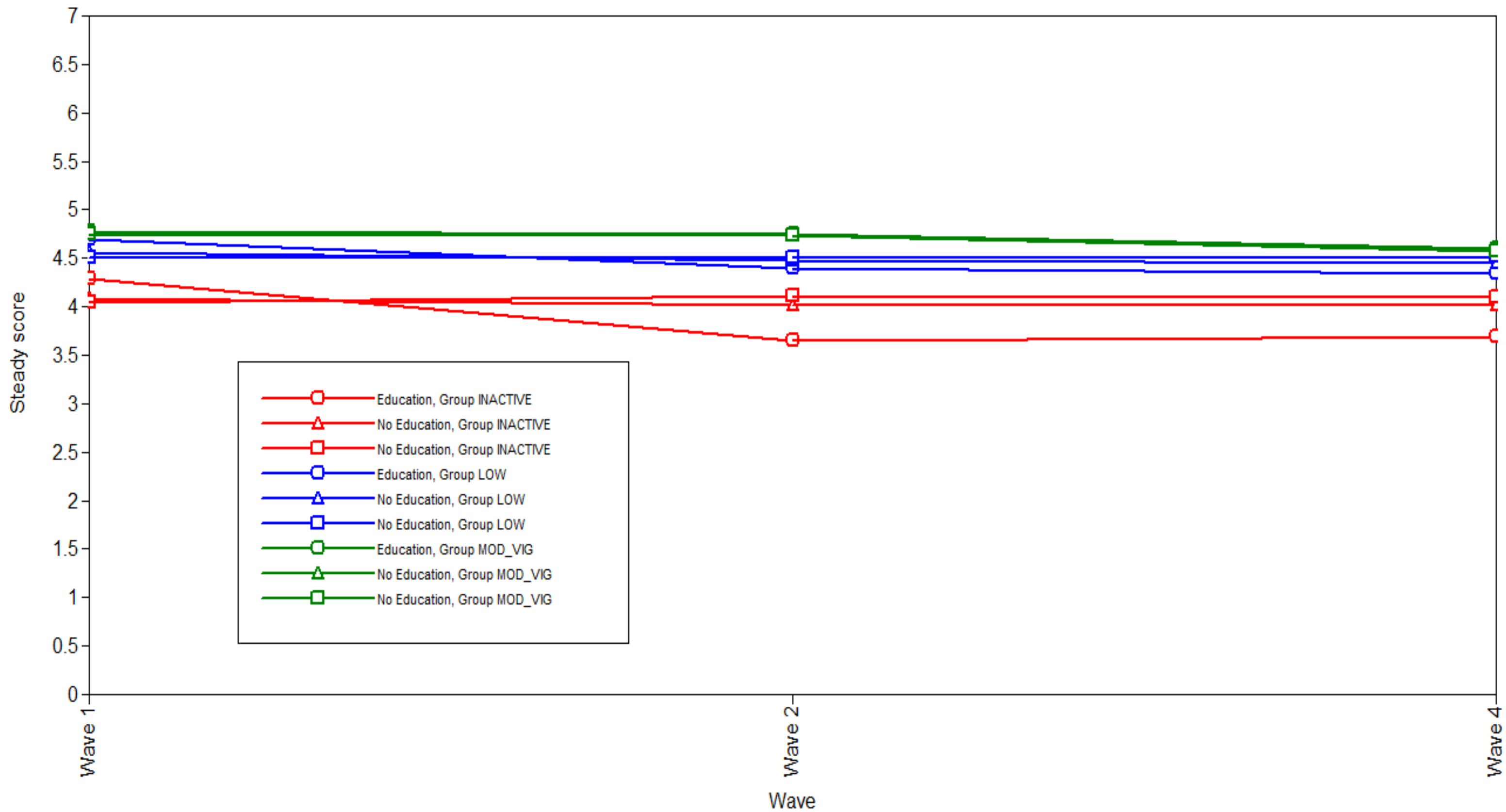


Figure 5.11
The effect of education level on the measure of steadiness for inactive, low, and moderate/vigorous PA groups

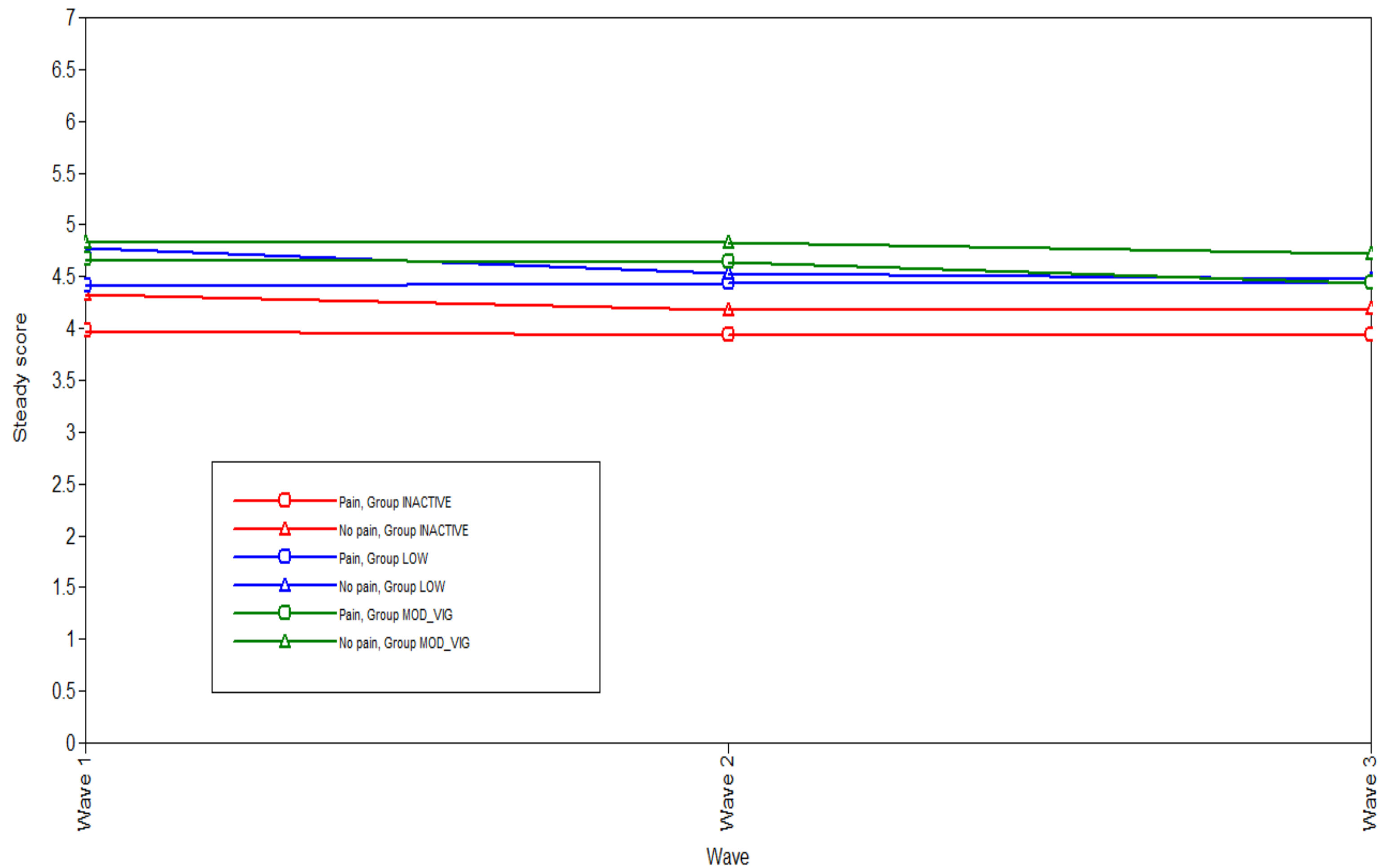


Figure 5.12
The effect of pain on the measure of steadiness for inactive, low, and moderate/vigorous PA groups

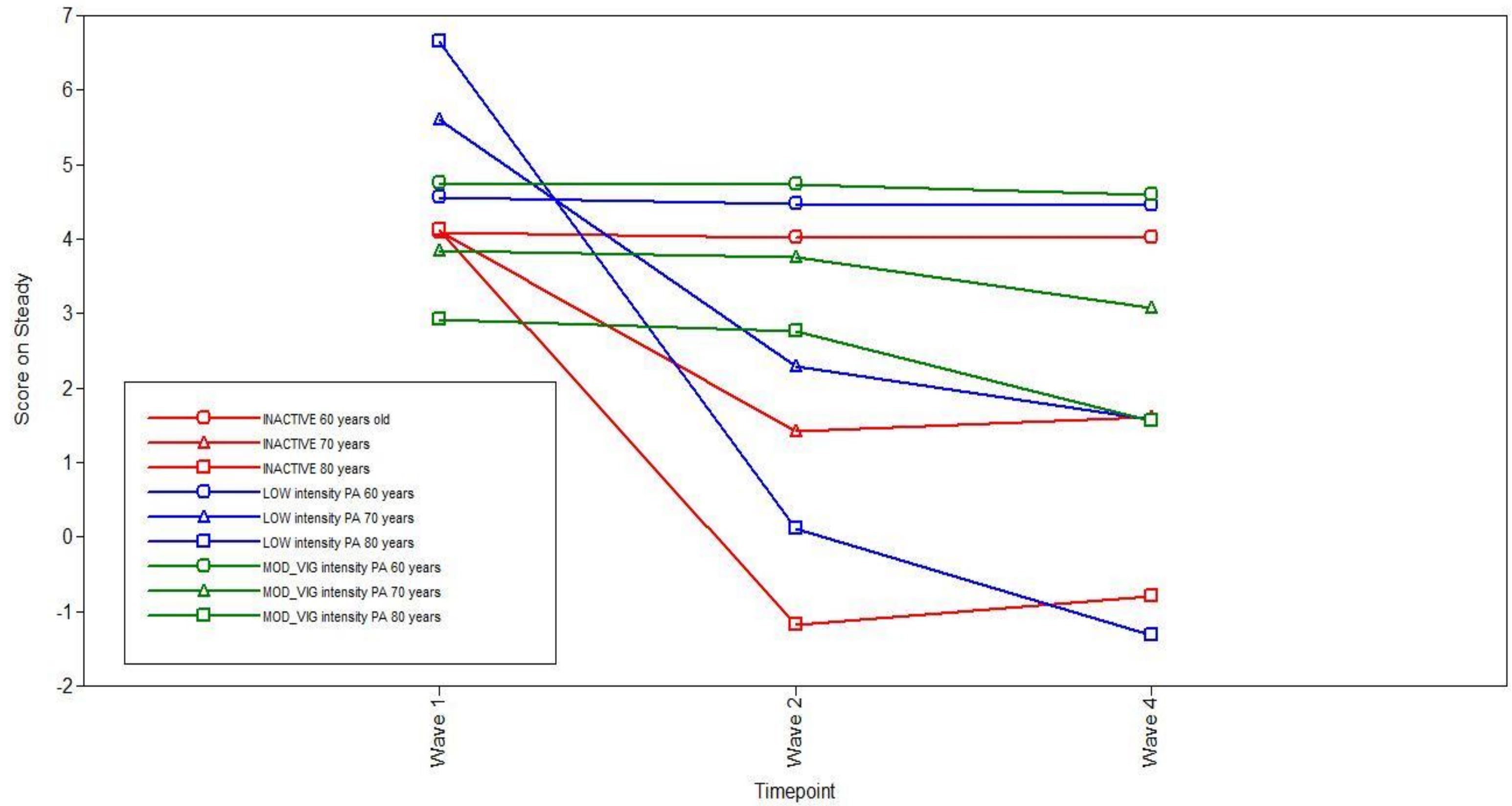


Figure 5.13
The trajectory of change in inactive, low, and moderate/vigorous PA groups on the measure of steadiness controlling for all variables by age

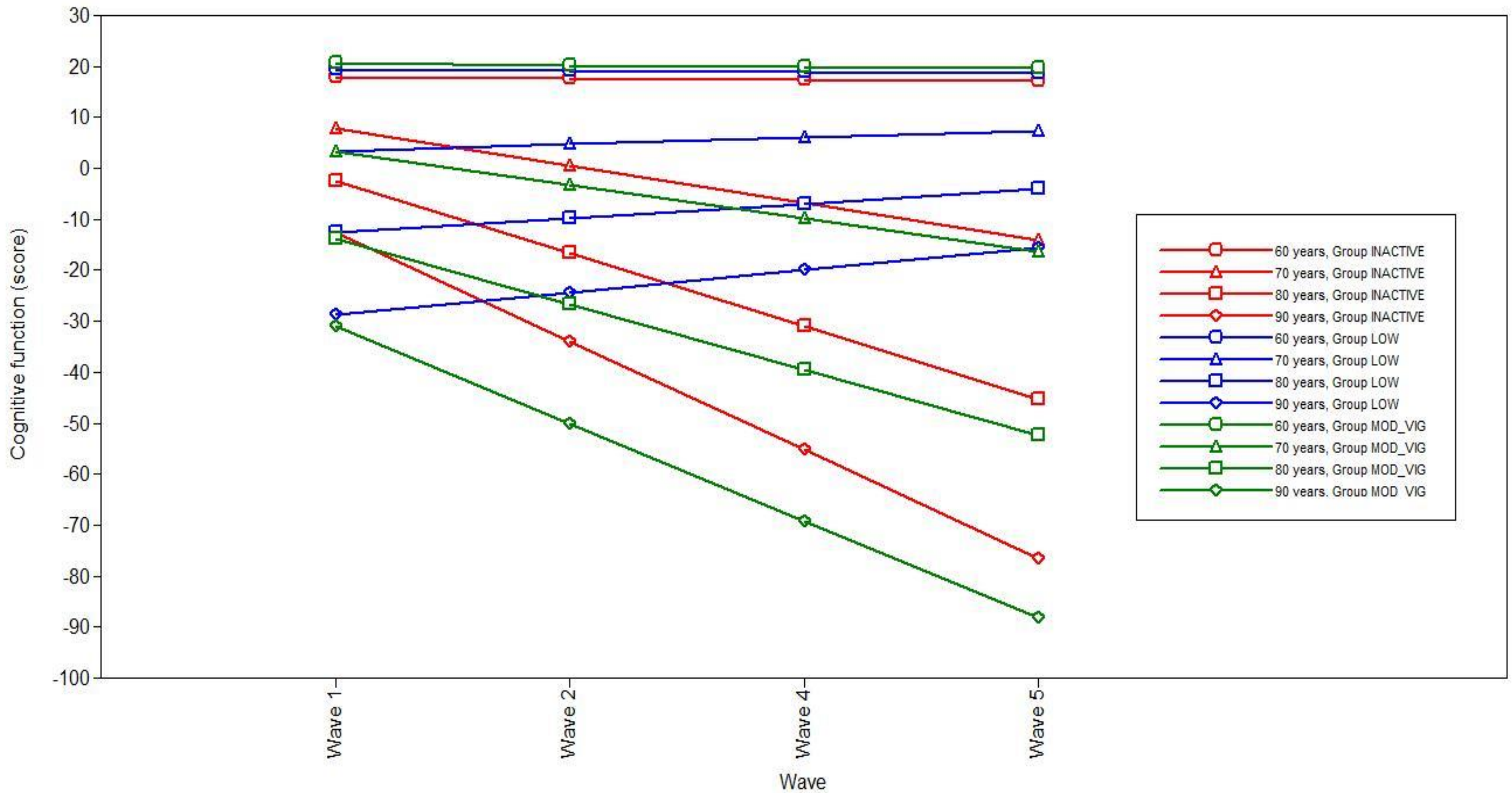


Figure: 5.14
The trajectory of change in inactive, low, and moderate/vigorous PA groups on the measure of cognitive function controlling for all variables by age

Chapter Six:

Validating the findings from TILDA, and ELSA analyses using the Northern Irish Cohort of Longitudinal Ageing (NICOLA)

6.1 Abstract

Background

Longitudinal studies such as TILDA, ELSA and NICOLA use standardised methods to increase the comparability of their findings. However, whilst the measures used may be similar across studies, differences that extend beyond the sample differences may impact on the validity and generalisability of the findings. Research suggests that establishing measurement invariance (MI) of the measures increases the opportunity to compare and generalise findings across studies. Therefore, this study explores the invariance of the measures of balance identified across TILDA, ELSA, and NICOLA studies to understand if the findings from one study can be generalised to another.

Methods

An investigation of the similarity between measures of balance across TILDA, ELSA, and NICOLA studies was carried out. Four similar measures of balance: a self-reported measure of vision, a self-reported measure of steadiness, and two objective measures of handgrip strength for the dominant hand were identified. A multi-group confirmatory factor analysis (MG-CFA) was used to test the invariance of the measures across TILDA, ELSA, and NICOLA studies to establish that the same measures are being used in each study. Firstly, the factor loadings for each measure were restricted to be equal across studies. Then, the intercepts of each measure across studies were restricted to be equal.

Model fit was evaluated using a Root Mean Square Error of Approximation (RMSEA) ≤0.05 with an upper limit (90% CI) ≤0.08; a Comparative Fit Index (CFI) ≥0.95; a Tucker Lewis Index (TLI) ≥0.95; and a Standardised Root Mean Square Residual (SRMR) ≤0.08 and where the levels of fit indices were not achieved, the modification indices were examined, and where appropriate, adjustments made. The correlations between measures across countries was also evaluated to understand if these measures provided a measure of balance.

Results

In total 29,048 participants were included in analyses of which 8,504 were from TILDA, 12,092 from ELSA, and 8,452 from NICOLA. Means and variances were different across studies for all four measures. Controls for covariates such as age and sex were not included. Intercepts and factor loadings were restricted to be equal across studies and fit statistics showed that the model described the data well, where RMSEA was 0.01 with an upper limit (90% CI) 0.02; a CFI 0.99; a TLI 0.99; and SRMR 0.02. Fit statistics also showed a low chi-squared value of 12.50 (7 degrees of freedom) indicating a good model fit, where chi-squared contribution from TILDA was 3.56, ELSA was 0.92, and NICOLA was 7.99. Correlations were low but statistically significant for steadiness with vision across TILDA (Est=0.80) and ELSA (Est=0.19) studies but not for NICOLA (Est=0.10). Correlations were statistically significant across all studies for grip strength with vision and grip strength with steadiness.

Conclusion

MG-CFA showed that measures used across studies of ageing such as TILDA, ELSA and NICOLA are invariant and therefore the findings relating to balance are validated and can be generalised across studies. Furthermore, correlations showed that the measures provided a measure of the composite measure of balance.

6.2 Introduction

Longitudinal studies such as TILDA (Chapter three, section 3.4.1), ELSA (Chapter four, section 4.4.1) and the Northern Irish Cohort of Longitudinal Ageing (NICOLA) study (described later in section 6.4.1) adhere to the Gateway to Global Ageing Data initiative (GGAD) (https://g2aging.org/). GGAD adherence is important because it provides a central repository of comparable questions and identically defined variables across surveys. Therefore, by adhering to GGAD, opportunities for the comparability of results across different countries or studies is possible. However, whilst the measures used in TILDA, ELSA, and NICOLA may be similar (Appendix XVIII), measurement invariance (MI) is also an important consideration for the interpretation and generalisation of findings across studies, and for the generalisability of the findings to other areas in the UK (Borsboom, 2006; Wirtz and Nachreimer, 2010).

For example, establishing MI involves running a series of constrained structural equation models to establish whether the factor loadings are equivalent (configural, metric, and scalar invariance, Chapter three). The establishment of MI means that differences in findings can be attributed to the differences in the population sample (e.g. age or sex) (Borsboom, 2006), rather than differences that extend beyond the population sample (Cheung & Rensvold, 2002; Hirschfeld & von Brachel, 2014; Van de Schoot et al., 2012; Wirtz and Nachreimer, 2010; Xu & Tracey, 2017). For example, where objective measures such as handgrip test have been used, differences may be caused using different dynamometers, or protocols for data collection. Additionally, where self-reported measures are used then these measures may be subject to biases such as recall, interpretation or social desirability (El-Gasim et al., 2013; Hassan, 2006; Mazor et al., 2002; Ramkissoona & Cole, 2011; Sakurai et al., 2013). Therefore, establishing MI of the measures used in the TILDA, ELSA, and NICOLA studies provides an opportunity to generalise and validate the findings from the model of free-living PA and balance (highlighted in Chapter three using data form the TILDA study) (Borsboom, 2006; Van de Schoot et al., 2012). It can help to identify whether the same measures are being used across studies and whether these measures have a relationship with the composite measure of balance (described in Chapter three, section 3.4.3). Multi-group confirmatory factor analysis (MG-CFA) within a SEM framework (as previously described in Chapter three, section 3.5, and Chapter five, section 5.5) has been widely used in educational psychology, social sciences, and medicine to explore MI across different groups (Chen, 2008; Wirtz and Nachreimer, 2010) and so will be used in this study to test for invariance.

Therefore, this study explores the invariance of the measures of balance used to establish the predictive model of free-living PA and balance in older adults (≥50 years) to establish whether the findings can be generalised across studies (Arlinghaus et al., 2012; Beran & Violato, 2010; Hoyle, 1995; Wirtz & Nachreiner, 2010).

6.3 Aims and objectives

The overall aim of this study is to validate the findings from the TILDA (Chapter three) and ELSA (Chapter five) analyses by using the indirect measures of balance from the NICOLA study. The following objectives were identified:

1) To identify the balance measures that are similar across TILDA, ELSA and NICOLA studies.

2) To validate that these measures are the same measures across TILDA, ELSA, and NICOLA datasets using a MG-CFA analysis to establish measurement invariance of the indirect measures of balance.

6.4 Methods

6.4.1 Study design

TILDA is described in Chapter three (section 3.4.1), and ELSA is described in Chapter four (section 4.4.1). NICOLA was launched in 2014 with 8,500 community-dwelling participants (≥50 years) randomly selected from the population in Northern Ireland (https://www.qub.ac.uk/sites/NICOLA/). The first wave of results was released at the end of 2017 containing both interview and health assessment data. The interview data captures information on health and social care utilisation; health behaviours; medication; mental, physical and cognitive health; employment; finances; retirement; social connectedness; social participation; driving and travel; housing; consumption and expectations for the future. The health assessment, carried out in a hospital setting, includes a review of cardiovascular, cognitive and respiratory function; physical activity; visual health; and body composition. Participants are also asked to provide biological samples for detailed laboratory analysis, including genetic analysis. It is expected that NICOLA will be followed-up for a period of at least 10 years and ethical approval for NICOLA was granted by the Queens University Research Ethics Committee.

6.4.2 Measures

Table 6.1 (page cx) identifies the similarities of balance measures used across all three studies (Appendix XVIII). Only 4 measures were similar across all three studies and these measures were measures of sensory and neuromuscular body systems only.

6.4.2.1 Sensory measures across TILDA, ELSA and NICOLA

All three studies use a similar validated self-reported measure of vision where TILDA and ELSA ask, ‘Is your eyesight poor; quite poor; average; good; excellent?’ and NICOLA asks, ‘At the present time would you say your eyesight is poor; quite poor; average; good; excellent?’ (El-Gasim et al., 2013; Zimdars, Nazroo & Gjonça, 2011).

6.4.2.2 Neuromuscular measures across TILDA, ELSA and NICOLA

All three studies use a similar self-reported measure of steadiness where TILDA and NICOLA ask, ‘While walking do you feel very unsteady; unsteady; quite steady; steady; very steady?’ and ELSA asks, ‘Do you have problems keeping balance while walking on a level surface always; often; sometimes; occasionally; never (Clark et al., 2005; Lindenberger et al., 2003).

All three studies include two measures of handgrip strength of the dominant hand using a dynamometer (Kg) (Bohannon & Shaubert, 2005). TILDA and NICOLA used a Baseline (Fabrication Enterprises Inc, White Plains, NY) hydraulic hand dynamometer. ELSA used a Smedley handheld dynamometer (Stoelting, Illinois, USA).

6.5 Statistical analysis

MG-CFA (Chen, 2008; Wirtz and Nachreimer, 2010) was used to test the invariance of the balance measures across TILDA, ELSA, and NICOLA studies in Mplus (version 7.4; Muthen & Muthen, Los Angeles, CA), because by establishing MI of the measures used provides an opportunity to generalise and validate the findings from TILDA (Chapter three), and ELSA analyses (Chapter five)

A complex analysis option in Mplus was used to account for clustering and stratification. Missing data were assumed to be missing at random (Schafer & Graham, 2002). A Maximum Likelihood Estimation (MLE) with robust standard errors was used to estimate the model, and is robust to non-normality (Enders, 2013; Yaun & Bentler, 2000).

Where there were single items for comparison across datasets then an empirical reliability of one is assumed where the observed variable is assumed to be identical to the unobserved construct. Where there are multiple items then a TAU-equivalent model was used where factor loadings are restricted to be equal and variances are unrestricted.

Firstly, factor loadings for each measure are constrained to be equal across studies to understand whether the same measure is being used across studies (metric invariance). Then, the intercepts of each measure across studies are constrained to be equal (scalar invariance) to allow the means to be compared to understand if differences are due to true change in the underlying construct. Finally, the equality of the correlations across countries are explored to understand if the measures represent the construct of balance as illustrated (Chapter three).

As we were only exploring the measures of balance in a cross-study comparison covariates such as sex, age were not included in the analysis. Model fit was evaluated using a Root Mean Square Error of Approximation (RMSEA) ≤0.05 with an upper limit (90% CI) ≤0.08; a Comparative Fit Index (CFI) ≥0.95; a Tucker Lewis Index (TLI) ≥0.95; and a Standardised Root Mean Square Residual (SRMR) ≤0.08 (Hoyle, 1995). Where the levels of fit indices were not achieved, the modification indices were examined, and where appropriate, adjustments made.

6.6 Results

Table 6.1 (page cx) shows that of a total of 21 measures of balance used across all three studies, four measures were similar across all three studies, six measures were similar across two studies (of which four were similar across NICOLA and ELSA; one was similar across NICOLA and TILDA, and one was similar across ELSA and TILDA), whilst 11 were not similar with any other study. The four measures identified measured neuromuscular and sensory body systems, but no measure was identified for cognition across all three studies.

Table 6.2 (page cxi) shows the descriptive statistics for those balance measures available across TILDA, ELSA, and NICOLA studies. In total 29,048 participants were included in analyses of which 8,504 were from TILDA, 12,092 from ELSA, and 8,452 from NICOLA. The proportion of data present in the analysis across all measures was high for TILDA (<70%), lower for ELSA (<61%) and less for NICOLA (36-59%) indicating a high proportion of the data was not obtained for NICOLA.

The final model of the analysis across all three studies is shown in Appendices XIX and XX. Tables 6.3 and 6.4 show the mean value and the variance of the mean for vision was lower for NICOLA (2.86/0.77) than for TILDA (3.65/0.88) or ELSA (3.44/1.04). The mean values and variances were higher for NICOLA for both measures of handgrip test (84.76/9613.31; 89.04/52236.49) than TILDA or ELSA. ELSA showed the highest mean for steadiness, but NICOLA showed the highest variance of the mean score (86.20). Intercepts and factor loadings were restricted to be equal across studies (configural and metric analysis). Table 6.5 (page cxiii) shows that the model described the data well, where RMSEA was 0.01 with an upper limit (90% CI) 0.02; a CFI 0.99; a TLI 0.99; and a SRMR 0.02. Fit statistics also showed a low chi-squared value of 12.50 (7 degrees of freedom) indicating a good model fit, where chi-squared contribution from TILDA was 3.56, ELSA was 0.92, and NICOLA was 7.99.

Table 6.3 (page cxii) shows the unstandardised results for the means (where TILDA is the reference study), for the latent construct of eye (one measure) was less for both ELSA and NICOLA than TILDA, whilst the means for the latent construct of steady (one measure) and grip (two measures) were greater for both ELSA (0.88; 0.27) and NICOLA (0.07; 0.27) than TILDA.

Table 6.4 (page cxii) shows that variances of the latent constructs of eye are similar across all three studies; is high for steady for NICOLA (86.20) than TILDA or ELSA; and is high across all three studies for grip.

Standardised results for correlations (Table 6.6, page cxiv) shows that correlations between the latent constructs of eye, steady and grip, whilst low, are significant except for steady with eye in the NICOLA study.

6.7 Discussion

6.7.1 Principal findings

Two similar neuromuscular measures and one similar sensory measure of balance were identified across TILDA, ELSA and NICOLA. A cognitive measure of balance was not identified. MG-CFA showed that the four measures of balance (vision, steadiness, and handgrip strength) are similar across TILDA, ELSA, and NICOLA studies, where the restrictions placed on the factor loadings and intercepts held across the three studies. Additionally, whilst the correlations between measures across studies was low, the results were still significant, and show that there is a consistently low correlation between measures across studies, therefore indicating the consistency of the measures used across studies. As a result, the finding from the TILDA and ELSA analyses can be validated and generalised (Borsboom, 2006; Van de Schoot et al., 2012; Wirtz and Nachreimer, 2010).

6.7.2 Strengths and weaknesses

Whilst the preparation and comparison of the three data sets took a considerable time to complete, a key strength of this analysis is that it uses data form three independent longitudinal studies of ageing, TILDA, ELSA, and NICOLA to understand the similarities of the measures. Additionally, the analysis uses MLR, thus allowing all participant data to be used, so reducing bias of the findings (Enders, 2013). Furthermore, this study includes both self-reported (vision and steadiness) and objective (handgrip test) measures to understand whether there were any differences caused by using different methods of data collection. However, a measure of cognition was not included in the analyses as there was no similar measure for cognitive function across all three studies.

Also, whilst this study included measures across multiple systems required for balance (e.g. neuromuscular and sensory body systems), a key weakness is that despite adhering to GGAD initiative there were few (four) indirect measures of balance that were similar across studies, and a cognitive measure of balance was not included.

6.7.3 Future research and clinical implications

The means and variances between studies across measures were varied, but controls for covariates such as age and sex were not included in the analysis as only the MI of the balance measures was explored. Therefore, future research should identify the sample differences between studies and control for these to explore the effects on measures of balance across studies. The implication of this finding is that the model of balance used in the TILDA and ELSA analyses can be applied to the NICOLA study, especially those using self-reported measure of vision, and objective measures of handgrip strength. Future studies should explore the relationship between handgrip and vision using the NICOLA data to understand the negative relationship between vision and strength where poorer eyesight indicates better grip strength not found in TILDA or ELSA studies.

6.8 Conclusion

The results from MG-CFA show that the model describes the data well when the factor loadings and intercepts of the observed measures of vision, steadiness and handgrip strength (test one and two) were restricted to be equal. Therefore, the findings from the TILDA analysis can be generalised to older adults across the UK and Ireland.

Table 6.1
Comparison of the indirect measures of balance across neuromuscular, cognitive and sensory body systems for TILDA, ELSA and NICOLA studies

Measure	TILDA	ELSA	NICOLA	Comparison
Sensory balance measures				
1. Is your eyesight?	Yes	Yes	Yes	3
2. Is your eyesight at a distance?	No	Yes	Yes	2
3. Is your eyesight close-up?	No	Yes	Yes	2
4. Is your hearing?	Yes	Yes	No	2
5. Do you find it difficult to hear with background noise?	No	Yes	No	n/a
6. Can you use a normal telephone?	Yes	No	No	n/a
Answer: 1=poor to 5=excellent				
Neuromuscular balance measures				
1. Steadiness when walking?	Yes	Yes	Yes	3
2. Steadiness when standing?	(summed)	No	Yes	n/a
3. Steadiness getting up?		No	Yes	n/a
4. Grip score dominant hand (kg)		Yes	Yes	3
5. Grip score dominant hand (kg)	Yes	Yes	Yes	3
6. Grip score dominant hand (kg)	Yes	Yes	No	n/a
7. TUG (secs)	No	No	Yes	2
8. Chair rise test (secs)	Yes	Yes	No	n/a
9. Gait speed test (secs)	No	Yes	No	n/a
Answer: 1=poor to 5=excellent	No			
Cognitive balance measures:				
1. MMSE	Yes	No	No	n/a
2. Date recall	No	Yes	No	n/a
3. Word recall (immediate)	No	Yes	Yes	2
4. Number of animals recalled	No	Yes	Yes	2
5. Prospective memory	No	Yes	No	n/a
6. Word recall (delay)	No	Yes	No	n/a
Answer: score				

Note: 3 - identifies similar measures across all three studies; 2 - identifies similar measures across two studies only; n/a - no comparison possible

Table 6.2
Descriptive statistics for similar measures of Vision, Steadiness, and Handgrip (2 tests) identified from TILDA, ELSA, and NICOLA studies

	Sample size			Mean (STD)			Skewness/Kurtosis		
Measure	**TILDA**	**ELSA**	**NICOLA**	**TILDA**	**ELSA**	**NICOLA**	**TILDA**	**ELSA**	**NICOLA**
Vision	8504	12090	5021	3.65 (0.88)	3.45 (1.04)	2.86 (0.77)	-0.33/-0.22	-0.17/-0.27	2.34/16.57
Steadiness	8504	11753	8452	3.75 (6.83)	4.64 (0.79)	4.36 (86.20)	33.57/215.79	-2.72/7.67	9.98/98.68
Handgrip test (test 1) (Kg)	6062	7505	3632	26.05 (97.91)	28.88 (130.83)	84.76 (9613.31)	0.52/-0.07	0.40/-0.30	3.85/12.87
Handgrip test (test 2) (Kg)	6069	7471	3632	25.22 (89.40)	29.42 (132.95)	89.04 (52236.49)	0.54/-0.15	0.38/-0.33	3.72/11.88

Table 6.3

Means for latent constructs of Eye (one measure), Steady (one measure), and Grip (two measures) between ELSA and NICOLA compared to TILDA (reference)

Variables	ELSA	NICOLA
Eye*	-0.21 (0.01)	-0.89 (0.04)
Steady*	0.88 (0.04)	0.07 (0.01)
Grip* (Kg)	0.27 (0.02)	0.27 (0.01)

Note: *TILDA is the reference group
Parameter estimates shown with standard errors in brackets**
Negative score indicates poorer eyesight

Table 6.4

Variances of the latent constructs of Eye, Steady and Grip across TILDA, ELSA and NICOLA

Latent construct	TILDA	ELSA	NICOLA
Eye	0.88	1.04	0.78
Steady	6.83	0.79	86.20
Grip	86.06	129.02	49065.51

Table 6.5
Fit statistics for the comparison of measures between TILDA, ELSA and NICOLA studies

Models	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	p-value	Estimate	90 % C.I.	CFI	TLI	Value
1.5 Intercepts & factor loadings are equal	466635.13	466908.26	533.48	9	0.0000	0.08	(0.07, 0.08)	0.89	0.79	0.01
1.6 Intercept on Grip removed	466196.25	466494.21	6.99	6	0.3220	0.04	(0.00, 0.01)	1.00	0.99	0.00
1.7 Test of correlations with restrictions	466264.94	466513.24	83.11	12	0.0000	0.03	(0.02, 0.03)	0.99	0.98	0.04
1.8 Correlations for Grip with eye removed across studies	466262.95	466527.80	85.95	10	0.0000	0.03	(0.02, 0.03)	0.99	0.97	0.03
1.5 Correlations of steady with grip; steady with eye removed across all studies	466207.28	466488.69	22.76	8	0.0037	0.01	(0.01, 0,02)	0.99	0.99	0.02
1.6 Correlations of steady with grip; steady with eye removed across all studies as well as steady with grip for TILDA only	466198.13	466487.81	12.49	7	0.0855	0.01	(0.00, 0.02)	0.99	0.99	0.02

Table 6.6
Correlations between latent constructs across TILDA, ELSA and NICOLA studies

Correlations	TILDA	ELSA	NICOLA
Steady with Eye	0.08 (0.01)	0.20 (0.01)	0.10 (0.06)*
Grip with eye	0.08 (0.01)	0.17 (0.01)	-0.07 (0.02)**
Grip with steady	0.05 (0.02)	0.23 (0.01)	0.0 (0.00)

Note
*Not statistically significant (p>0.05)
Parameter estimate shown with standard error in brackets
** Grip with Eye is negative for NICOLA indicating if you have strong grip then you have poorer eyesight is not found in TILDA or ELSA.

Chapter Seven:

Discussion

7.1 Introduction

Evidence supports the proposition that balance is critical for health and well-being in an ageing population, but the ageing process itself is a key risk factor for poor balance and accidental falls which can result in disability and death in older adults (≥50 years). PA guidelines and fall prevention policy recommend MVPA for general health benefits and fall prevention. However, guidance does not appear to be relevant to the population at which it is aimed as older adults are more likely to be engaged in PA that is not MVPA. Furthermore, guidance for fall prevention in older adults at low risk of falling is lacking, and evidence supporting the benefits of PA such as LPA lacks methodological robustness. Consequently, the overall aim of this thesis was to contribute to our understanding of the relationship between PA and balance in older adults (≥50 years).

As described in Chapter one (section 1.6), the thesis comprised of two strands. The purpose of the first strand (Chapter two) was to help define and frame the research topic by understanding the existing literature with the aim of identifying the characteristics of PA and its relationship with balance, as well as identify gaps in the research in a systematic way. The purpose of the second strand (Chapters three-six) was to develop a robust predictive model of PA and balance using the understanding gained from the systematic review (Chapter two).

7.2 Summary of findings

Chapter one provided an overview of the importance of fall prevention in older adults and highlighted the gaps in our understanding of the relationship between PA and balance in older adults (≥50 years). It identified a lack of appropriate guidelines for older adults for fall prevention; a lack of understanding of the effects of LPA on balance; a lack of methodologically robust studies; and the challenges of measuring PA and balance.

Chapter two systematically reviewed the existing literature (30 papers; n=1,574 participants) relating to the association between PA and balance in older adults (≥50 years) at lower risk of falling. The focus was on free-living PA that included PA for the purposes of leisure, occupation, and travel. The methodological quality and adequacy of the measures included were assessed. The results suggested that more active older adults who engaged in free-living PA between one to 21 years' duration (e.g. tai chi, dance, walking and cycling) compared with less active older adults experienced better balance. However, most of the evidence was from cross-sectional studies (25/26 studies) of moderate methodological quality, and a much smaller proportion from RCTs (four studies) of low methodological quality. For example, cross-sectional studies cannot explore the effects of PA over time or individual differences; sample size across all studies was small (20-170 participants) giving rise to Type II errors; self-reported measures of PA were used across studies thus limiting the conclusions drawn, and balance assessment was not comprehensive with only one study including measures across all three body systems for balance. Most studies included neuromuscular measures (19 of 30 studies), less included cognitive measures (ten of 30 studies), and even fewer included sensory measures (three of 30 studies). Subsequently, the review concluded that whilst free-living PA is beneficial to balance measures, there are key areas that must be addressed to improve the quality of the evidence.

Consequently, Chapter three explored the association between PA and balance in older adults (≥50 years) using a longitudinal analysis of data from the TILDA study over a two-year period. This study addressed the methodological challenges of existing evidence relating to small sample size, and duration of study. It also included multiple measures of balance therefore measure balance as a multidimensional construct and addressed measurement error. Using CFA within a SEM framework the analyses identified that multiple indirect functional measures of balance such as MMSE (cognitive); TUG, handgrip test, and steadiness (neuromuscular); and vision and hearing (sensory) collectively provided a composite measure of balance (Horak, 1995; Sibley et al., 2015). Additionally, the results demonstrated that balance declined with age (Horak, 2006), but that PA, the activity of everyday living, can improve or maintain balance over a two-year period. Furthermore, the findings also suggested older adults who are female (WHO, 2007); used medication (Blake, 1988); had a problem performing an ADL (Bandeen-Roche et al., 2015; Tak et al., 2013) resulted in lower activity levels which adversely affected balance. Additionally, increased age, fear of falling (Bandeen-Roche et al., 2015), lower education (Chen et al., 2015; Preston, et al., 1998); pain (Marmot et al., 2002); and high alcohol consumption (WHO, 2007) were also confirmed as important risk factors for poor balance. The findings did not support that sleep quality or history of falls were significant for either PA or balance.

The TILDA analysis included only two waves of data, and whilst the analysis indicated the amount of change, it could not provide a trajectory or rate of change over time (Duncan & Duncan, 2009). As a result, Chapters four and five used data from the ELSA study of ageing which includes six waves of data over a 10-year period. Firstly, Chapter four explored the most appropriate method of developing a composite measure of PA using a self-reported categorical measure of PA collected by the ELSA study. It compared two conventional scoring approaches (Demakakos et al., 2010; Hamer et al., 2014) with that of an LCA approach using three latent classes. The results showed that the level of distinction between subgroups across all three approaches was greater using an LCA of three classes.

Chapter five explored the trajectory of change in balance by level of PA intensity using the LCA PA measure which consisted of inactive, LPA, and MVPA groups. Firstly, CFA was used to identify the measures with the strongest relationship with the construct of balance. Then, using an LGM approach, the individual change and differences over time were explored (Duncan & Duncan, 2009). The findings of the CFA supported that multiple indirect functional measures of balance, such as measures of cognition (prospective memory, fluency), measures of neuromuscular (chair rise test, gait speed test, handgrip test, and steadiness), and sensory system (vision and hearing) collectively provided a model for balance assessment in an older population (Horak, 1995; Sibley et al., 2015). The self-reported measure of steadiness, and the objective measures of cognitive function and gait speed were identified as having the strongest relationship with the composite measure of balance. LGM analysis then showed that there was a dose-response relationship between PA intensity and balance performance in older adults ≤70 years. where MVPA increases gait speed, cognitive function, and self-rated steadiness, and slows the rate of decline across all age groups. Additionally, LPA also has a role to play where older adults ≥70 years had better gait speed and cognitive function than those in MVPA groups. Therefore, LPA whilst not as beneficial as MVPA, may still have benefits for those adults who may not be physically capable of achieving MVPA levels.

Lastly, research suggests that establishing measurement invariance across independent studies increases comparability of results and the opportunity to generalise findings (Arlinghaus et al., 2012; Beran & Violato, 2010; Hoyle, 1995; Wirtz & Nachreiner, 2010). Therefore, Chapter six explored the invariance of four similar measures of balance that measured neuromuscular and sensory body system, identified across TILDA, ELSA, and NICOLA studies. There was no similar measure of cognitive function available across all three studies. The results of MG-CFA showed that the measures identified were associated with the construct of balance, and were also invariant across TILDA, ELSA and NICOLA studies. Therefore, the findings relating to balance can be generalised across studies.

The objective of the current chapter is to further discuss some of the key points raised from the previous chapters and concludes by highlighting the potential contribution the overall findings have to clinical practice, as well as the implications for future research.

7.3 Methodological design

Reiner et al.'s (2013) systematic review of longitudinal studies (n = 15 studies, 288,724 participants aged 18-85 years) exploring the long-term effects of PA on non-communicable diseases (NCD) such as coronary heart disease and diabetes, highlights the importance of studying change over time to fully understand prevention and ongoing effect. Similarly, research supports that balance declines with age (Chodzko-Zajko et al., 2009), and PA such as LPA may require study over time to understand the associated benefits (Bauman et al., 2016; Hoffmann et al., 2016; Morris & Hardman, 1997). However, exploring the effect of a behaviour such as PA on balance over time is challenging.

The Medical Research Council (MRC) (2008) framework for complex interventions (www.mrc.ac.uk/complexinterventionsguidance) recognises the importance of long-term follow-up to determine whether the outcomes predicted occur or whether short-term changes are maintained. Despite these guidelines, clinical trials that explore the longer-term effects of complex interventions such as PA are lacking (Craig et al., 2008). For example, the systematic review carried out as part of this thesis (Chapter two) identified only four RCTs, with the longest period of study being six months; none included any follow-up. This finding is further echoed by Howe et al.'s (2011) systematic review where the length of trial ranged from one to twelve months with the most frequent being three months; 75 of the 94 included studies did not have any follow-up, and 19 studies that did include a follow-up only had a follow-up duration of between six weeks to one year.

Fortunately, RCTs are beginning to emerge such as Iliffe et al.'s (2015) (n=1,254 participants) RCT exploring structured and planned programmes of exercise that included a walking component with a 12-month post intervention follow-up. However, high cost, and the practical difficulty in adhering to protocol over prolonged timeframes, continue to limit the ability to explore change over time (Smith et al., 2011). Additionally, even if interventions were implemented in the same way as the trials over a longer duration, the fidelity of the intervention may still be compromised, and similar outcomes may not be achieved (Cohen et al., 2015; Hoffmann et al., 2016).

As a result, whilst RCTs provide direction and guidance relating to research, real-life studies may complement clinical trials by offering additional information not only on the efficacy of the intervention, but also on the effectiveness in a much broader patient population. Therefore, longitudinal studies provide an opportunity to have a meaningful impact at a population level. For example, this thesis was limited in time and budget and so it used secondary data analysis of longitudinal data from studies of ageing (TILDA, ELSA, and NICOLA). Whilst studies such as TILDA, ELSA and NICOLA were not specifically designed to address the research question under evaluation within this thesis, and the robustness of the findings are influenced by the quality and design of the data (Cooke & Iwashyna, 2013), the analysis provides an opportunity to explore the relationship between PA and balance over time (2, 10, and 1 year respectively) using repeated measures from large population samples (8,504; 10,601; and 8,500 participants respectively). As a result, it explores PA and balance in a less expensive and time intensive way, with an ablility to make recommendations that are ecologically valid (Cooke & Iwashyna, 2013; Smith et al., 2011).

Thus, whilst both design approaches have limitations from both a practical and theoretical perspective, perhaps an optimal approach to study design for the exploration of complex behaviours such as PA and its influence on balance is to use RCTs as a way of directing research, and then longitudinal or secondary data analysis of longitudinal studies to demonstrate the effectiveness of PA in a real-life setting.

7.4 Valid and reliable measures

An important component of research design is the use of valid and reliable measures to increase the robustness and generalisability of the findings (Kimberlin & Winterstein, 2008). For example, the Framingham study (MacMahon et al., 1990) shows how addressing measurement error increases the robustness of the findings. McMahon et al. (1990) explored the association between diastolic blood pressure (DBP) and coronary heart disease (CHD) in nine observational studies (n=420,000 participants with a mean follow-up of 10 years) but identified that studies had ignored measurement error caused by the diluting effects of random fluctuations in DBP, and so these studies had underestimated the association between DBP and CHD by 60%. Whilst MacMahon et al.'s (1990) study highlights the importance of minimising measurement error for clinical practice, it also highlights that it is a concept that is rarely addressed (Borsboom, 2006; Coggon et al., 2003; Nachtigall et al, 2003). Consequently, this study sought to improve the reliability and validity of the measures of PA and balance. It used SEM to account for measurement error which enabled the separation of random variance due to population sample differences such as culture, sex or age from variance due to other factors including error in measurement (McCoach et al., 2007).

For example, as highlighted in Chapter one (section 1.3.1) self-reported measures of PA are more widely used in longitudinal studies that include older adults such as TILDA and ELSA (Chapters three, four, and five) because they can measure behavioural change (Shephard, 2003), and are relatively easy to implement and interpret (Bauman et al., 2009; Dishman et al., 2001; Welk, 2002). However, self-reported PA may be influenced by health status, mood, depression, anxiety, or cognitive ability (Murphy, 2009), seasonal variation in PA patterns, social desirability, or recall issues (Dyrstad et al., 2014; Saelens et al., 2012). These biases may result in an overestimation of the actual levels of activity, where vigorous intensity activity may be overestimated, and low to moderate intensity activity underestimated, and so ways of minimising this bias is necessary to improve the robustness of the results as well as increasing the opportunity to compare and generalise results (Borsboom, 2006; Hagenaars & McCutcheon, 2002; McCoach et al., 2007; Van de Schoot et al., 2012). The results from Chapter four addressed this issue showing that an LCA approach incorporating BCH to address measurement error, was a more appropriate measure than conventional scoring approaches. An LCA approach provided a greater distinction or demarcation between subgroups and identified patterns of activity in complex data that were ignored by conventional scoring approaches.

Additionally, developing the model of balance (Chapters three and five) involved using multiple proxy measures of balance that included objective and self-reported measures, but guidance on which combination of individual measures are more effective is lacking (Horak et al., 2006; Sibley et al., 2015; Vieira et al., 2016). The task of identifying combinations of measures that provide adequate measurement is made more complex by the number of different indirect measures available. For example, the systematic review carried out as part of this study (McMullan et al., 2018) identified 23 different indirect or proxy measures of across studies (n=30 studies) which supports the findings from previous systematic reviews that identified 25 (Howe et al., 2011), and 66 (Sibley et al., 2015) different indirect balance measures. Furthermore, the results from the systematic review found that only one observational study (Fong & Ng, 2006) included indirect functional measures that assessed all the systems required for balance performance (neuromuscular, cognitive and sensory).

The use of CFA within a SEM approach enabled the mapping of multiple indirect measures of balance to establish whether the indirect measures provided by TILDA and ELSA comprehensively measured balance (Bollen, 1989; McCoach et al., 2007; Raykov & Marcoulides, 2000). For example, TILDA analyses (Chapter three) showed objective cognitive measures such as MMSE; objective neuromuscular measures such as TUG, handgrip test, and a self-reported measure of steadiness; and sensory measures such as self-reported vision and hearing have a statistically significant relationship with balance over a two-year period. The objective measure of TUG (Est=0.71, S.E.=0.04), and the self-reported measure of steadiness (Est=0.86, S.E.=0.04) had the strongest relationship with the construct of balance. Additionally, ELSA analyses (Chapter five) showed that five measures of cognitive function; three objective neuromuscular measures, gait speed, handgrip strength, chair rise test, a measure of self-reported steadiness, and sensory measures of self-reported vision and hearing have a statistically significant relationship with balance across a 10-year period. Three neuromuscular measures that included objective measures of chair stand test (Est=1, S.E.=0.02) and gait (Est=0.78, S.E.=0.03), the self-reported measure of steadiness (Est=0.60, S.E.=0.02), as well as cognitive measures (Est=0.61, S.E.=0.01) showed the strongest relationship with balance. Both analyses showed that the sensory measures of self-reported vision and hearing had a weaker relationship with balance (TILDA: vision Est=0.27, S.E.=0.02, hearing Est=0.23, S.E.=0.03; ELSA: Eye Est=0.41, S.E.=0.01; Ear Est=0.25, S.E.=0.02), and in both analyses a residual correlation was introduced between the variables across all waves because these measures showed a variance not explained by the construct of balance. This was perhaps due to a method effect as both sensory measures are self-reported and so may be open to recall bias (Hassan, 2006), interpretation bias (Mazor et al., 2002), or or a bias caused by over- or under-estimation (Ramkissoona & Cole, 2011; El-Gasim et al., 2013; Sakurai et al., 2013).

Furthermore, testing the measurement invariance of similar measures across TILDA, ELSA, and NICOLA (Chapter six) using a SEM approach enabled the validity of the measures to be tested. MG-CFA identified that the measures of self-reported steadiness, vision, and two measures of grip strength were the same across all three studies. Therefore, the findings can be generalised across the UK and Ireland.

Hence, SEM minimises measurement error and so increases the validity and reliability of the measures of both PA and balance. It also enables the theoretical construct of balance to be developed using multiple different observed measures. Consequently, the approach improves the robustness and generalisability of the findings from this research. Despite these advantages, SEM is a data analytic technique, and is still subject to the same limitations as other statistical methods relating to the quality of the study design or data. For example, SEM assumes that sample selection is randomised; it requires a large sample size; the use of modifications to improve the fit statistics is open to theoretical interpretation by the researcher; and causality is an assumption rather than a consequence of SEM (McCoach et al., 2007). As a result, the practical use of SEM for longitudinal studies that are not robustly designed may be limited, and its use may be more appropriate for secondary data analysis of longitudinal data from studies such as TILDA, ELSA and NICOLA that have been robustly designed.

7.5 Cumulative benefits of PA for balance in later life

The analysis of TILDA and ELSA showed that balance declined over two or a ten-year period. For example, TILDA analysis (Chapter three) showed that the mean of balance decreased by 25% over a two-year period, and ELSA analysis showed that balance declined over a ten-year period where the difference in the mean of balance at wave one and six showed a decline of 29%. This is supported by evidence that the body systems required for balance decline either through disease or degeneration beginning from 40 years of age (Chodzko-Zajko et al., 2009). This is further supported by the statistics highlighting that the rate of falls increases from 28-35% in adults ≤65 years to 32-42% in adults ≥70 years (WHO, 2007; 2015). Therefore, interventions and policy are required to prevent falls in the general population and PA promotion is needed throughout the lifecycle (Skelton & Mavroeidl, 2018).

Previous research suggests that PA is a modifiable factor that can influence balance in later life. For example, Cooper et al.'s (2011) study (n= 5,362 participants) exploring the effects of habitual PA of moderate to vigorous intensity (e.g. sports, exercise and moderate to vigorous leisure activities) over three time points in middle-aged adults (36-53 years) found cumulative benefits for physical performance. Additionally, Lang et al. (2007) found that regular exercise consistently improved sarcopenia, physical function, cognitive performance and mood in both frail and non-frail older adults. The findings from the systematic review (Chapter two) and analysis of TILDA data (Chapter three) extend these findings. In the systematic review we identified that free-living PA, that included PA for the purposes of leisure, occupation and travel,carried out over a duration of between one to 21 years improves measures of balance in healthy older adults at less risk of falling. Additionally, free-living PA carried out over a two-year period can improve comprehensive balance performance. An extra MET-minute of PA can improve balance performance by 4% in the short-term and has a cumulative effect on balance over a two-year period where an extra MET-minute of PA per week improved balance by 5%. Furthermore, PA level at baseline also has a cumulative effect on PA level over a two-year period (Est=0.28; SE=0.02).

This finding suggests that the older adults who have more active lifestyles earlier in life, are at less risk of falling because this has a protective effect on balance either via its effects on balance at wave one, or its effects via PA at wave two. Therefore, PA promotion and guidelines need to promote PA throughout the lifecycle to ensure effective fall prevention.

7.6 Implications for policy

PA guidelines focus on MPVA for health benefits in older adults (CMO, 2011; Kuehne & Brannan, 2018). However, despite these guidelines and the plethora of research supporting the benefits of MPVA (Gillespie et al., 2012; Howe et al., 2011; Milton et al., 2018) public health systems have failed to introduce successful strategies to support older adults to include this in their everyday living activity (Bann et al., 2015; Foster & Armstrong, 2018; Milton et al., 2018; Murtagh et al. 2015; Schutzer & Grave, 2004). In fact, research suggests that older adults are more likely to engage in LPA than MVPA in their daily lives (Arnardottir et al., 2013; Franco et al., 2015), and perhaps due to the focus of policy on MPVA, less is understood about the effects of LPA (Milton et al., 2018; Powell et al., 2011).

Analysis of ELSA data (Chapter five) suggested that MPVA is more beneficial than LPA in terms of the cumulative effects on balance over a ten-year period with the effects of MPVA showing a slower rate of decline in measures of grip test, self-rated steadiness, and cognitive function, and better results in those older adults ≤70 years supporting that PA and balance have a dose-response relationship (Powell et al., 2011). Research suggests that barriers to MVPA may be that it is perceived as activity that cannot be integrated into everyday living and research calls for guidance to include examples of activities that can support fall prevention that can be integrated into everyday living activity

(Foster & Armstrong, 2018). For example, barriers such as lack of time, lack of motivation or access to appropriate facilities, perceived risk to health, fear of falling, (Baert et al., 2011; Cavill & Foster, 2018; Chao et al., 2000; Franco et al., 2015; Schutzer & Graves, 2004) suggest that the concept of MVPA is considered as an additional 'chore' that sits outside of everyday activity. So, whilst it is recognised that MVPA can elicit better results in relation to balance performance, if guidelines are to be relevant to the people at which they are aimed then more detail regarding the types of activity that can help achieve these targets may be needed. For example, the Netherlands have emphasised the contribution of everyday activities such as stair climbing within their PA guidelines (Health Council of the Netherlands, 2017), and so perhaps guidelines need to be more explicit in terms of the actual activities that are beneficial.

Furthermore, older adults may be more physically capable of carrying out LPA than MVPA (Arnardottir et al., 2013; Franco et al., 2015) and the longitudinal analysis using the ELSA data (Chapter five) extends existing research by showing that LPA is beneficial to older adults ≥70 years, therefore supporting that LPA may be more beneficial to older adults at higher risk of falling (Izawa et al., 2017; Osuka et al., 2015; Varma et al., 2014). Additionally, the ELSA analysis indicated that whilst providing less benefits than MVPA, LPA still slows the decline in gait speed, steadiness, and cognitive function than being inactive (Powell et al., 2011). Therefore, whilst MVPA may be more beneficial for balance, LPA still has a cumulative effect on balance over time that is more beneficial to older adults than being inactive (Duvivier et al., 2013). As a result, the promotion of LPA for fall prevention in older adults ≥70 years should be a consideration for future guidelines.

In summary, a ‘one size fits all approach’ for effective fall prevention is inappropriate. Instead, fall prevention policy and healthcare promotion should include MVPA for older adults ≤70 years and LPA for older adults ≥70 years to address the variable health issues and barriers to PA in older adults.

7.7 Implications for clinical practice

This study presents a model of balance that can guide balance assessment within clinical practice because it uses multiple indirect balance measures therefore ensuring that balance is comprehensively measured. More robust screening measures can help to assess risk of falls more accurately, and target interventions more appropriately. It is recognised that further research exploring the practical application of combining measures is needed, but objective measures such as gait speed or TUG as well as the self-reported measure of steadiness (“when standing...” “when getting up from a chair...” and “when walking...do you feel?”, (1) very steady; (2) slightly steady; (3) slightly unsteady; (4) very unsteady) have been shown to have a strong relationship with balance, and so should at least be considered for integration into balance assessment as a minimum requirement.

In addition, the results show that free-living PA, that is not just exercise, can benefit balance in the immediate term as well as have a cumulative effect over time. Therefore, programmes of activity for older adults may be developed that may not only be beneficial to balance, but also more appropriate to this population given that this population are failing to meet exercise guidelines

(Hallal et al., 2012). Additionally, there are both immediate and cumulative benefits of PA on balance in older adults, thus increased activity should be promoted in older adults to ensure the maintenance or improvement in balance in later life. Furthermore, advice and guidance for fall prevention in older adults needs to be explicit and try to address the barriers to PA adherence where for example, everyday activities that are of MVPA are identified and whilst MVPA is promoted to older adults ≤70 years, LPA is promoted to older adults ≥70 years.

7.8 Implications for future research

Future research should consider exploring the use of the composite measure of balance proposed in this thesis within clinical practice to assess the practicality of implementation and refine the measure. Also, both self-reported and objective measures are included, and future research should consider both the convergent validity between indirect and direct measures to further assess the appropriateness of the model of balance. For example, an objective measure of vision such as LogMAR (Minimal Angle of Resolution) charts and hearing such as the pure tone audiometry test (PTA) to address the methods effect highlighted in section 5.7.3. Also, future research should include methodologically robust RCTs with longer durations. Furthermore, an investigation of the benefits of LPA using longitudinal analysis should be considered to provide more explicit guidelines relating to dose requirements for LPA.

7.9 Conclusion

This study has used longitudinal analysis of data addressing measurement error as well as testing the generalisability of the results to show that free-living PA carried out over one to 21 years improves measures of balance in older adults at lower risk of falling; that a multi-system approach including indirect functional measures provides a model for balance; that free-living PA has a cumulative effect on balance over time; that PA and balance have a dose-response relationship where higher intensity PA reduces the rate of decline in balance performance and leads to better balance performance in older adults ≤70 years; that LPA also elicits benefits to balance performance in older adults ≥70 years; and that the findings from the analysis of TILDA and ELSA can be generalised to other studies of ageing such as NICOLA. Therefore, the findings support that early participation in PA can benefit future balance performance, and higher intensity PA reduces the rate of decline in older age. Furthermore, the findings extend our understanding of the effects of LPA where LPA has benefits for balance in older adults ≥70 years in addition to the general health benefits and so should be promoted in healthcare. Therefore, as Rudyard Kipling (1902) pointed out *“the cure for this ill, is not to sit still, or frowst by a book by the fire, but to grab a large hoe and a shovel also and dig to you gently perspire”*.

References

Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, O'Brien WL, Bassett DR, Schmitz KA, Emplalncourt PO, Jacobs DR & Leon AS. (2000). Compendium of physical activities: an update of activity codes and MET intensities. Med. Sci. Sports Exerc., 32(9), S498-S516.

Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett Jr DR, Tudor-Locke C, Greer JL, Vezina J, Whitt-Glover MC, Leon AS. (2011). Compendium of physical activities: a second update of codes and MET values. Med Sci Sports Exerc, 43, 1575–81.

Akaike, H. (1987). Factor analysis and AIC. Psychometrika, 52, 317–332

Angevaren M, Vanhees L, Nooyens AC, Wendel-Vos CG, Verschuren WM. (2010). Physical activity and 5-year cognitive decline in the Doetinchem cohort study. Ann Epidemiol., 20(6), 473-479.

Aoyagi, Y Park, H, Watanabe, E, Park, S & Shephard, RJ. (2009). Habitual physical activity and physical fitness in older Japanese adults: the Nakanojo Study. Gerontology, 55(5): 523-531.

Arlinghaus A, Lombardi DA, Willetts JL, & Christiani DC. (2012). A structural equation modelling approach to fatigue-related risk factors for occupational injury. Am J Epidemiol, 176(7), 597-607.

Arnardottir NY, Koster A, Van Domelen DR et al. (2013). Objective measurements of daily physical activity patterns and sedentary behaviour in older adults: Age, Gene/Environment Susceptibility-Reykjavik Study. Age Ageing, 42, 222–229.

Asparouhov T & Muthén B. (2014). Auxiliary variables in mixture modelling: using the BCH method in mplus to estimate a distal outcome model and an arbitrary secondary model. Mplus Web notes, 21 (2), 1-22.

Baert V, Gorusa E, Metsa T, Geertsa C & Bautmansa I. (2011). Motivators and barriers for physical activity in the oldest old: A systematic review. Ageing Research Reviews, 10, 464– 74.

Bailey IL & Lovie JE. (1976.) New design principles for visual acuity letter charts. Am J Optom Physiol Opt. 53 (11), pp. 740–745.

Baker LD, Frank LL, Foster-Schubert K. Green PA, Wilkinson CW, McTiernan A, Cholerton BA, Plymate ST, Fishel MA, Watson S, Duncan GE, Mehta PD & Crafta S. (2010). Aerobic exercise improves cognition for older adults with glucose intolerance, a risk factor for Alzheimer’s disease. J Alzheimers Dis., 22(2), 569-579.

Bakk Z. & Vermunt JK. (2015). Robustness of stepwise latent class modelling with continuous distal outcomes, structural equation modelling: A Multidisciplinary Journal, Doi: 10.1080/10705511.2014.955104.

Bandeen-Roche K, Seplaki CL, Huang J, Buta B, Kalyani RR, Varadhan R, Xue Q, Walston JD & Kasper JD. (2015). Frailty in older adults: a nationally representative profile in the United States. J Gerontol A Biol Sci., 70(11), 1427-1434.

Bann D, Hire D, Manini T, Cooper R, Botoseneanu A, McDermott MM, . . . LIFE SG. (2015). Light intensity physical activity and sedentary behavior in relation to body mass index and grip strength in older adults: Cross-sectional findings from the lifestyle interventions and independence for elders (LIFE) study. PloS One, 10(2), 1. Doi: http://dx.doi.org/10.1371/journal.pone.0116058.

Bartholomew DJ, Steele F, Moustaki I & Galbraith J. (2008). Analysis of multivariate social science data (2nd edition). Boca Raton, FL: Chapman & Hall.

Bauman A, Ainsworth B, Bull F, Craig C, Hagströmer M, et al. (2009) Progress and Pitfalls in the Use of the International Physical Activity Questionnaire (IPAQ) for Adult Physical Activity Surveillance. Journal of Physical Activity and Health 6, S5–S8. PMID: 19998844.

Bauman MP, Hekler EB, Haskell WL et al. (2010). Objective light intensity physical activity associations with rated health in older adults. Am J Epidemiol, 172, 1155–1165.

Bauman, A., Merom, D., Bull, F., Buchner, D. & Fiatarone, M.A. (2016). Updating the evidence for physical activity: summative reviews of the epidemiological evidence, prevalence, and interventions to promote ‘active aging’. Gerontologist. 58, S268-S280. DOI: 10.1093/geront/gnw031.

Bauman A, Phongsavan P, Schoeppe S & Owen N. (2006). Physical activity measurement- a primer for health promotion. Promotion & Education, 13(2), 92-103. DOI:10.1177/10253823060130020103.

Bentein K, Vandenberg R, Vandenberghe C & Stinglhamber F. (2005). The role of change in the relationship between commitment and turnover: A latent growth modeling approach. Journal of Applied Psychology, 90 (3), 468-482. Doi: 10.1037/0021-9010.90.3.468.

Beran TN & Violato C. (2010). Structural equation modelling in medical research: a primer. BMC Research Notes, 3, 267. http://www.biomedcentral.com/1756-0500/3/267.

Best JR, Davis JC & Liu-Ambrose T. (2015). Longitudinal analysis of physical performance, functional status, physical activity, and mood in relation to executive function in older adults who fall. JAGS, 63(6), 1112-1120.

Black A, Wood J, Lovie-Kitchin J E & Newman B. (2008). Visual Impairment and postural sway among older adults with glaucoma. Optometry and vision science: official publication of the American Academy of Optometry. 85. 489-97. Doi:10.1097/OPX.0b013e31817882db.

Blake AJ, Morgan K, Bendall MJ, Dallosso H, Ebrahim SBJ, Arie THD, Fentem PH, & Bassey EJ. (1988). Falls by elderly people at home: prevalence and associated factors. Age Ageing 17, 365–372.

Blondell SJ, Hammersley-Mather R & Veerman JL (2014) Does physical activity prevent cognitive decline and dementia? A systematic review and meta-analysis of longitudinal studies. BMC Public Health, 14, 510. Doi: 10.1186/1471-2458-14-510.

Bohannon RW & Shaubert KL. (2005). Test–retest reliability of grip-strength measures obtained over a 12-week interval from community-dwelling elders. J Hand Ther., 18, 426–428.

Bolck A, Croon M, & Hagenaars J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. Political Analysis, 12, 3–27.

Bollen K. (1989). Structural equations with latent variables. New York, NY: John Wiley & Sons, Inc; 1989.

Bonato P. (2005). Advances in wearable technology and applications in physical medicine and rehabilitation. J. Neuroeng. Rehabil, 2 (2), 1-4. Doi:10.1186/1743-0003-2-2.

Borenstein M, Hedges LV, Higgins JPT & Rothstein HR. (2009). Introduction to meta-analysis. John Wiley & Sons Ltd. ISBN: 978-0-470-05724-7 chapter 40.

Borsboom, D. (2006). The attack of the psychometricians. Psychometrika, 71(3), 425–440. Doi: 10.1007/s11336-006-1447-6.

Brewer JK & Sindelar PT. (1998). Adequate sample size: a priori and post-hoc considerations. The Journal of Special Education, 21 (4), 74-84. Doi:10.1177/002246698802100409.

Brigola AG, Rossetti ES, Rodrigues dos Santos B, Neri AL, Zazzetta MS, Inouye K, & Lost Pavarini SC. (2015). Relationship between cognition and frailty in elderly: a systematic review. Dement Neuropsychol., 9(2), 110-119.

Brooke-Wavell, K & Cooling, VC. (2009). Fall risk factors in older female lawn bowls players and controls. Journal of Aging & Physical Activity, 17(1):123-130.

Buatois S, Gauchard GC. Aubry C, Benetos A, Perrin,P. (2007). Current Physical Activity Improves Balance Control during Sensory Conflicting Conditions in Older Adults. International Journal of Sports Medicine, 28(1): 53-58.

Bucknix F, Rolland Y, Reginster JY, Ricour C, Petermans J, & Bruyere O. (2015). Burden of frailty in the elderly population: perspectives for a public health challenge. Archives of Public Health, 73(19). Doi: 10.1186/s13690-015-0068-x.

Byrne BM, Shavelson RJ & Muthen B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. Psychological Bulletin, 3, 456-466.

Cadore EL, Rodriguez-Manas L, Sinclair A, & Izquierdo M. (2013). Effects of different exercise interventions on risk of falls, gait ability, and balance in physically frail older adults: a systematic review. Rejuvenation Research, 16(2): 105-114. Doi: 10.1089/rej.2012.1397.

Cavill NA & Foster CEM. (2018). Enabler and barriers to older people’s participation in strength and balance activities: a review of reviews. JFSF, 3(2), 105-113. Doi: 10.22540/JFSF-03-105.

Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. Journal of Classification, 13, 195–212.

Chao D, Foy CG & Farmer D. (2000). Exercise adherence among older adults: challenges and strategies. Control Clin Trials., 21, 212S-217S.

Chartered Society of Physiotherapy (CSP) (2018). Get up and go: a guide to staying steady. Downloaded on 6th June 2018 from: http://www.csp.org.uk/publications/get-go-guide-staying-steady

Chen FF. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. Journal of Personality and Social Psychology, 95(5), 1005.

Chen CM, Chang WC, & Lan TY. (2015). Identifying factors associated with changes in physical functioning in an older population. Geriatrics Gerontology, 15, 156-164.

Chen X, Mao G & Leng SX. (2014). Frailty syndrome: an overview. Clinical Interventions in Ageing., 9, 433-441.

Cheung GW & Rensvold RB. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. Structural Equation Modeling, 9, 233-255. Doi:10.1207/S15328007SEM0902_5.

Chia EM, Mitchell P, Rochtchina E, Foran S, Golding M, & Wang JJ. (2006). Association between vision and hearing impairments and their combined effects on quality of life. Arch Ophthalmol, 124, 1465-1470.

Chief Medical Officer (CMO). (2011). Start active, stay active. A report on physical activity for health from the four home countries' Chief Medical Officers. k11; Downloaded on 12th October 2015 from: https://www.gov.uk/government/publications/start-active-stay-active-a-report-on-physical-activity-from-the-four-home-countries-chief-medical-officers.

Chodzko-Zajko WJ, Proctor DN, Fiatarone Singh MA, Minson CT, Nigg CR, Salem GJ, & Skinner JS. (2009). American college of sports medicine stand on exercise and physical activity for older adults. Medicine & Science in Sports & Exercise, 41(7), 1510-1530.

Clark, D.O., Callahan, C.M., & Counsell, S.R. (2005). Reliability and validity of a steadiness score. Journal of the American Geriatrics Society, 53 (9), 1582-1586.

Clark DB, Jones BL, Wood DS, & Cornelius JR. (2006). Substance use disorder trajectory classes: diachronic integration of onset age, severity, and course. Addict Behav. 31, 995–1009.

Coggon D, Rose GDM & Barker DJP. (2003). Epidemiology for the uninitiated, fourth edition. Chapter 4. Measurement error and bias. Downloaded on 31st August 2018 from: https://www.bmj.com/about-bmj/resources-readers/publications/epidemiology-uninitiated/4-measurement-error-and-bias.

Cohen A, Goto S, Schreiber K, et al. (2015). Why do we need observational studies of everyday patients in the real-life setting? Eur Heart J Suppl. 17(suppl D), D2-D8.

Cooke CR & Iwashyna TJ. (2013). Using existing data to address important clinical questions in critical care. Crit Care Med., 41 (3), 886-896. Doi: 10.1097/CCM.0b013e31827bfc3c.

Cooper R, Kuh D, Hardy R, Mortality Review Group, FALCon & HALCyon Study Teams. (2010). Objectively measured physical capability levels and mortality: systematic review and meta-analysis. BMJ, 341, c4467. Doi: 10.1136/bmj.c4467.

Cooper R, Mishra GD, & Kuh D. (2011). Physical activity across adulthood and physical performance in midlife. Findings from a British birth cohort. Am J Prev Med., 41(4), 376-384.

Craig P, Cooper C, Gunnell D, Haw S, Lawson K, Macintyre S, Ogilvie D, Petticrew M, Reeves B, Sutton M & Tomson S. (2012). Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. J Epidemiol Community Health. 66, 1182–1186. Doi: 10.1136/jech-2011-200375.

Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I & Petticrew M. (2008). Developing and evaluating complex interventions: the new Medical Research

Council guidance. The BMJ, 337, a1655. Doi: http://doi.org/10.1136/bmj.a1655.

Craig CL, Marshall AL, Sjöström M, Bauman A, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, Sallis JF & Oja P. (2003). International Physical Activity Questionnaire: 12-country reliability and validity. Medicine and Science in Sports and Exercise. 1381-1395.

Cumming RG, Salkeld G, Thomas M, Szonyi G. (2000). Prospective study of the impact of fear of falling on activities of daily living, SF-36 scores, and nursing home admission. J Gerontol A Biol Sci Med Sci, 55: M299-305. Doi: 10.1093/gerona/55.5.M299.

Curran PJ & Willoughby MJ. (2003). Reconciling theoretical and statistical models of developmental processes. Development and Psychopathology, 15, 581-612.

Demakakos P, Hamer M, Stamatakis E & Steptoe A. (2010). Low-intensity physical activity is associated with reduced risk of incident type 2 diabetes in older adults: evidence from the English Longitudinal Study of Ageing. Diabetologia, 53, 1877–1885. Doi: 10.1007/s00125-010-1785-x.

Department of Health (DoH). (2011). Start Active, Stay Active: A report on physical activity from the four home countries' Chief Medical Officers. Downloaded on 7th June 2018 from:

https://www.gov.uk/government/publications/start-active-stay-active-a-report-on-physical-activity-from-the-four-home-countries-chief-medical-officers.

Dewhurst, S., Nelson, N., Dougall, P.K. and Bampouras, T.M. (2014). Scottish country dance: benefits to functional ability in older women. Journal of Aging & Physical Activity, 22(1): 146-153.

DiPietro L, Caspersen CJ, Ostfeld AM & Nadel ER. (1993). A survey for assessing physical activity among older adults. Medicine and Science in Sport and Exercise. 25, 628-642.

Dishman RK, Waskburn RA & Schoellar DA. (2001). Measurement of physical activity. QUEST, 53, 295-309.

Doherty A, Jackson D, Hammerla N, PloÈ T, Olivier P, Granat MH, White T, van Hees VT, Trenell MI, Owen CG, Preece SJ, Gillions R, Sheard S, Peakman T, Brage S, Wareham NJ. (2017). Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. Study. PLoS ONE 12(2), e0169649. Doi: 10.1371/journal.pone.0169649.

Dohrn IM, Kwak L, Oja P, Sjostrom M & Hagstromer M. (2018). Replacing sedentary time with physical activity: a 15 year follow up of mortality in a national cohort. Clinical Epidemiology, 10, 179-186. Doi: https://doi.org/10.2147/CLEP.S151613.

Duncan TE & Duncan SC. (2009). The ABC's of LGM: An introductory guide to latent variable growth curve modelling. Soc Personal Psychol Compass, 3(6), 979-991. Doi: 10.1111/j1751-9004.2009.00224.x.

Duvivier BMFM, Schaper NC, Bremers MA, van Crombrugge G, Menheere PPCA. (2013). Minimal intensity physical activity (standing and walking) of longer duration improves insulin action and plasma lipids more than shorter periods of moderate to vigorous exercise (cycling) in sedentary subjects when energy expenditure is comparable. PLoS ONE 8(2): e55542. Doi:10.1371/journal.pone.0055542

Dyrstad SM, Hansen DM, Holme IM & Anderssen SA. (2014). Comparison of self-reported versus accelerometer-measured physical activity. Med Sci Sports Exec, 46, 99-106.

El-Gasim, M., Munoz, B., West, S.K., & Scott, A.W. (2013). Associations between self-rated vision score, vision tests, and self-reported visual function in the Salisbury eye evaluation study. Invest Ophtbalmol Vis Sci., 54, 6439-6445.

Enders, C.K. (2013). Dealing with missing data in developmental research. Child Dev Perspect., 7, 27–31.

Fedak KM, Bernal A, Capshaw ZA & Gross S. (2015). Applying the Bradford Hill criteria in the 21st century: how data integration has changed causal inference in molecular epidemiology. Emerg Themes Epidemiol,12(14), 1-9. Doi: 10.1186/s12982-015-0037-4.

Ferrer E, Hamagami F & McArdle JJ. (2009). Modeling latent growth curves with incomplete data using different types of structural equation modeling and multilevel software. Structural Equation Modeling: A Multidisciplinary Journal, 11(3), 452-483, Doi: 10.1207/s15328007sem1103_8.

Folstein, MF, Folstein, SF, & McHugh, PR. (1975). 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res: 12: 189-198.

Fong, S. & Ng, G.Y. (2006). The effects on sensorimotor performance and balance with Tai Chi training. Archives of Physical Medicine & Rehabilitation, 87(1): 82-87 6p.

Fong SS. Ng SS. Liu KP. Pang MY. Lee HW. Chung JW. Lam PL. Guo X. Manzaneque,J.M. (2014). Musculoskeletal strength, balance performance, and self-efficacy in elderly Ving Tsun Chinese martial art practitioners: implications for fall prevention. Evidence-Based Complementary and Alternative Medicine, Vol. 2014, Article ID 402314: 6 pages: http://dx.doi.org/10.1155/2014/402314.

Forbes SC, Forbes D, Forbes S, Blake CM, Chong LY, Thiessen EJ, Rutjes AWS, & Little JP. (2015). Exercise interventions for maintaining cognitive

function in cognitively healthy people in late life. Cochrane Database of Systematic Reviews, Issue 5. Art. No.: CD011704.
Doi: 10.1002/14651858.CD011704.

Foster C & Armstrong MEG. (2018). What types of physical activities are effective in developing muscle and bone strength and balance? JFSF, 3(2), 58-65. Doi: 10.22540/JFSF-03-058.

Franco MR, Tong A, Howard K, Sherrington C, Ferreira PH, Pinto RZ & Ferreira ML. (2015). Older people's perspectives on participation in physical activity: a systematic review and thematic synthesis of qualitative literature. Br J Sports Med, 1-9. Doi: 10.1136/bjsports-2014-094015.

Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M & Bennell KL. (2015). Exercise for osteoarthritis of the knee. Cochrane Database Syst Rev, Issue 1. Art. No.: CD004376. Doi: 10.1002/14651858.CD004376.pub31.

Gao KL, Hui-Chan CW & Tsang WW. (2011). Golfers have better balance control and confidence than healthy controls. European Journal of Applied Physiology, 111(11), 2805-2812.

Gateway to Global Ageing Data (GGAD). https://g2aging.org/

Gauchard, G.C., Jeandel, C., Tessier, A. and Perrin, P.P. (1999). Beneficial effect of proprioceptive physical activities on balance control in elderly human subjects. Neuroscience Letters, 273(2): 81-84.

Gauchard, G.C., Jeandel, C. and Perrin, P.P. (2001). Physical and sporting activities improve vestibular afferent usage and balance in elderly human subjects. Gerontology, 47(5): 263-270.

Gauchard, G.C., Gangloff, P., Jeandel, C. and Perrin, P.P. (2003). Influence of regular proprioceptive and bioenergetic physical activities on balance control in elderly women. Journals of Gerontology Series A-Biological Sciences & Medical Sciences, 58(9): M846-50.

Guadagnin, E.C., da Rocha, E.S., Mota, C.B. and Carpes, F.P. (2015). Effects of regular exercise and dual tasking on spatial and temporal parameters of obstacle negotiation in elderly women. Gait & Posture, 42(3): 251-256.

Gillespie LD, Robertson MC, Gillespie WJ, Sherrington C, Gates S, Clemson LM, & Lamb SE. (2012). Interventions for preventing falls in older people living in the community. Cochrane Database of Systematic Reviews, 9. Art. No.: CD007146. Doi: 10.1002/14651858.CD007146.pub3.

Graham JE, Ostir GV, Fisher SR, & Ottenbacher KJ. (2008). Assessing walking speed in clinical research: a systematic review. Journal of Evaluation in Clinical Practice ISSN 1356-1294.

Gyllensten, A.L., Hui-Chan, C.W.Y. and Tsang, W.W.N. (2010). Stability limits, single-leg jump, and body awareness in older Tai Chi practitioners. Archives of Physical Medicine & Rehabilitation, 91(2): 215-220 6p.

Hagenaars J A & McCutcheon AL. (2002). Applied latent class analysis. Cambridge: Cambridge University Press. Downloaded on 5th June 2018 from: http://catdir.loc.gov/catdir/samples/cam031/2001037649.pdf

Hakim, R.M., DiCicco, J., Burke, J., Hoy, T. and Roberts, E. (2004). Differences in balance related measures among older adults participating in Tai Chi, structured exercise, or no exercise. Journal of Geriatric Physical Therapy, 27(1): 11-15 5p.

Hakim, R.M., Kotroba, E., Cours, J., Teel, S. and Leininger, P.M. (2010). A cross-sectional study of balance-related measures with older adults who participated in tai chi, yoga, or no exercise. Physical & Occupational Therapy in Geriatrics, 28(1): 63-74.

Hallal PC, Andersen LB, Bull F, Guthold R, Haskell W & Ekelund U. (2012). Global physical activity levels: surveillance, progress, pitfalls, and prospects. The Lancet, 380, 247-257.

Hamer M, Lavoie KL, & Bacon SL. (2014). Taking up physical activity in later life and healthy ageing: the English longitudinal study of ageing. Br J Sports Med, 48, 239–243. Doi:10.1136/bjsports-2013-092993.

Harada ND, Chiu V, King AC & Stewart AL. (2001). An evaluation of three self-report physical activity instruments for older adults. Medicine and Science in Sports and Exercise, 962-970.

Hartling L, Milnea A, Hamma MP, Vandermeera B, Ansari M, Tsertsvadzec A, & Drydena DM. (2013). Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. Journal of Clinical Epidemiology, 66: 982-993.

Haskell WL & Kiernan M. (2000). Methodologic issues in measuring physical activity and physical fitness when evaluating the role of dietary supplements for physically active people. The American Journal of clinical Nutrition, 72 (2), 541S-550S. https://doi.org/10.1093/ajcn/72.2.541S.

Haslam D. (2015). Prevention better than cure, says NICE Chair. Address to the Royal College of General Practitioners (RCGP) conference in Glasgow, downloaded on 1st August 2018 from: https://www.nice.org.uk/news/article/prevention-better-than-cure-says-nice-chair.

Hassan, E. (2006). Recall bias can be a threat to retrospective and prospective research designs. The Internet Journal of Epidemiology, 3(2), 1-7. 3.4. 10.5580/2732.

Health Council of the Netherlands (2017). Physical activity guidelines 2017. Publication no. 2017/08e. Downloaded on 8th September 2018 from: https://www.gezondheidsraad.nl/en/task-and-procedure/areas-of-activity/preventie/dutch-physical-activity-guidelines-2017

Health Survey England (HSE) (2017). Health survey for England 2016: Physical activity in adults. Health and Social Care Information Centre. www.digital.nhs.uk. Downloaded on 18/01/18.

Herzog R, Álvarez-Pasquin MJ, Díaz C, Del Barrio JL, Estrada JM, & Gil A. (2013) Are healthcare workers' intentions to vaccinate related to their knowledge, beliefs and attitudes? a systematic review. BMC public health, 13: 1–17.

Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savović J, Schulz KF, Weeks L, Sterne JAC & Cochrane bias methods group. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ, 343, d5928. Doi: 10.1136/bmj.d5928.

Higgins, JPT & Green, S. (editors) (2011). Cochrane handbook for systematic reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration. Available from www.cochrane-handbook.org.

Hillsdon, MM, Brunner, EJ, Guralnik, JM, & Marmot, MG. (2005). Prospective study of physical activity and physical function in early old age. Am J Prev Med., 28: 245-250.

Hirji KF & Fagerland MW. (2009). Outcome based subgroup analysis: A neglected concern. Trials, 10 (33). Doi: 10.1186/1745-6215-10-33·

Hirschfeld G & von Brachel R. (2014). Improving Multiple-Group confirmatory factor analysis in R – A tutorial in measurement invariance with continuous and ordinal indicators. Practical Assessment, Research & Evaluation, 19(7), Downloaded on 26th July 2018 from: http://pareonline.net/getvn.asp?v=19&n=7.

Hoffmann TC, Maher CG, Briffa T, Sherrington C, Bennell K, Alison J, Singh MF & Glasziou PP. (2016). Prescribing exercise interventions for patients with chronic conditions. CMAJ, 1-9. Doi:10.1503/emaj.150684.

Horak, FB. (1995). Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls? Age and Ageing, 35-52. Doi: 10/1093/ageing/afl077.

Horak FB. (2006). Postural orientation and equilibrium: what do we need to know about neural control of balance to prevent falls? Age and Ageing, 35-52. Doi: 10/1093/ageing/afl077.

Horak FB & Macpherson JM. (1996). Postural orientation and equilibrium. In: Rowell LB, Shepard JT, eds. Handbook of Physiology: Section 12, Exercise Regulation and Integration of Multiple Systems. New York: Oxford University Press, 255–292.

Horak FB, Wrisley DM, & Frank J. (2009). The Balance Evaluation Systems Test (BESTest) to Differentiate Balance Deficits. Physical Therapy. 89(5), 484-498. Doi:10.2522/ptj.20080071.

Howe TE, Rochester L, Neil F, Skelton DA, & Ballinger C. (2011). Exercise for improving balance in older people. Cochrane Database of Systematic Reviews. Issue 11. Art. No. CD004963.
Doi: 10.1002/14651858.CD004963.pub.

Howley TE. (2001). Types of activity: resistance, aerobic and leisure versus occupational physical activity. Med. Sci. Sports Exerc, 33(6), S364-S369.

Hox J & Stoel RD. (2005). Multilevel and SEM approaches to growth curve modelling. Encyclopaedia of Statistics in Behavioural Sciences. 3, 1296-1305.

Hoyle, R.H. (1995). Structural Equation Modelling: Concepts, issues and applications. Thousand Oaks, CA: Sage Publications, 1-132.

Idler EL & Angel RJ. (1990). Self-rated health and mortality in the NHANES-I epidemiologic follow-up study. Am J Public Health, 80: 446–52.

Iliffe S, Kendrick D, Morris R, Griffin M, Haworth D, Carpenter H, Masud T, Skelton DA, Dinan-Young S, Bowling A & Gage H. (2015). Promoting physical activity in older people in general practice: ProAct65+ cluster randomised controlled trial. Br J Gen Pract, e731-e738, Doi: 10.3399/bjgp15X687361.

Izawa KP, Shibata A, Ishii K, Miyawaki R, Oka K. (2017). Associations of low-intensity light physical activity with physical performance in community-dwelling elderly Japanese: A cross-sectional study. PLoS ONE, 12(6), e0178654. https://doi.org/10.1371/journal.pone.0178654.

Jacobs R, Goddard M & Smith PC. (2005). How robust are hospital ranks based on composite performance measures? Medical Care, 43 (12), 1177-1184.

Jacobs JV, Van Tran K, & Nutt JG. (2006). An alternative clinical postural stability test for patients with Parkinson's disease. J Neurol, 253(11), 1404-1413.

Jiang L, Chen S, Zhang B, Beals J, Mitchell CM, Manson SM, Roubideaux Y, & the special diabetes program for Indians diabetes prevention demonstration project. (2015). Longitudinal patterns of stages of change for exercise and lifestyle intervention outcomes: An application of latent class analysis with distal outcomes. Prev Sci, 17, 398-409. Doi: 10.1007/s11121-015-0599-y.

Joint Health Surveys Unit (2007). Health Survey for England Physical Activity Validation Study: substantive report. Information Centre for Health and Social Care, Leeds, UK. 2007.

Jung T & Wickrama KAS. (2008). An introduction to latent class growth analysis and growth mixture modelling. Social and Personality Psychology Compass 2/1, 302-317. Doi: 10.1111/j.1751-9004.2007.00054. x.

Kaplan GA & Camacho T. (1983). Perceived health and mortality: a nine-year follow-up of the human population laboratory cohort. Am J Epidemiol ,117: 292–304.

Karlsson MK, Magnusson H, Schewelov T & Rosengren BE. (2013). Prevention of falls in the elderly-a review. Osteoporosis International, 24, 747-762. Doi: 1007/s00198-012-2256-7.

Kearney, P.M., Cronin, H., & O’Regan, C. (2011) Cohort profile: The Irish Longitudinal Study of Ageing. Int J Epidemiol, 40, 877-884.

Kimberlin CI & Winterstein AG. (2008). Validity and reliability of measurement instruments used in research. Am J Health-Syst Pharm, 65, 2276-2284.

Kongstead A & Nielsen AM. (2017). Latent class analysis in health research. Journal of Physiotherapy, 63, 55-58.

Kuehne JD & Brannan MGT. (2018). Only the strong: why we need more focus on strengthening and balance activities in physical activity. JFSF, 3(2), 56-57. Doi: 10.22540/JFSF-03-056.

Lacey RJ, Strauss W, Rathod T. (2015). Clustering of pain and its associations with health in people aged 50 years and older: cross-sectional results from the North Staffordshire Osteoarthritis Project. BMJ Open, 5: e008389. Doi: 10.1136/bmjopen-2015-008389.

Lamb SE, Jorstad-Stein EC, Hauer K, & Becker C. (2005). Development of a common outcome data set for fall injury prevention trials: The prevention of falls network Europe consensus. JAGS, 53:1618–1622.

Lamb SE, Becker C, Gillespie LD, Smith JL, Finnegan S, Potter R, et al. (2011). Reporting of complex interventions in clinical trials: development of a taxonomy to classify and describe fall-prevention interventions. Trials, 12, 125-132. Doi: 10.1186/1745-6215-12-125.

Landi F, Abbatecola AM, Provinciali M et al. (2010). Biogerontology, 11, 537. Doi: https://doi.org/10.1007/s10522-010-9296-1.

Lang IA, Guralnik JM & Melzer D (2007). Physical activity in middle-aged adults reduces risks of functional impairment independent of its effect on weight. JAGS 55 (11), 1836–1841.

Lanza ST & Rhoades BL. (2013). Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. Prev Sci., 14(2), 157-168. Doi: 10.1007/s11121-011-0201-1.

Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT & the Lancet Physical Activity Series Working Group. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet, 380, 219–229.

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, & Moher D. (2009). The PRISMA statement for reporting systematic reviews and meta-analysis of studies that evaluate healthcare interventions: explanations and elaboration. BMJ. 339: b2700. PMID: 19622552. DOI: 10.1136/bmj.b2700.

Lienhard, K; Schneider, D; & Maffiuletti, NA. (2013). Validity of the Optogait photoelectric system for the assessment of spatiotemporal gait parameters. Medical Engineering & Physics, 35: 500-504.

Lin FR & Ferrucci L. (2011). Hearing Loss and Falls Among Older Adults in the United States. Arch Intern Med; 172(4): 369–371. Doi:10.1001/archinternmed.2011.728.

Lin CC & Whitney S. (2012). Quantification of static and dynamic balance while maintaining and changing body position. Top. Geriatr. Rehabil., 28, 17–26.

Lindenberger EC, Landefeld CS, Sands LP, Counsell SR, Fortinsky RH, Palmer RM, Kresevic DM, & Covinsky KE. (2003). Unsteadiness reported by older hospitalized patients predicts functional decline. J Am Geriatr Soc, 51: 621-622.

Lo, Y, Mendell, N, & Rubin D. (2001). Testing the number of components in a normal mixture. Biometrika, 88, 767–778.

Lord, SR, Wars, JA, Williams, P & Anstey, KJ. (1994). Physiological factors associated with falls in older community dwelling women. J Am Geriatr Soc. 42: 1110-1117.

Lu, X., Siu, K.C., Fu, S.N., Hui-Chan, C.W. and Tsang, W.W. (2013). Tai Chi practitioners have better postural control and selective attention in stepping down with and without a concurrent auditory response task. European Journal of Applied Physiology, 113(8): 1939-1945.

MacMahon S, Peto R, Cutler J, Collins R, Sorlie P, Neaton J, Abbott R, Godwin J, Dyer A & Stamler J. (1990). Blood pressure, stroke and coronary heart disease: Part 1, prolonged in differences in blood pressure: prospective observational studies corrected for the regression dilution bias. Lancet, 335, 765-774.

Magidson J & Vermunt J. (2004). Latent class models. In D. Kaplan (Ed.), Handbook of quantitative methodology for the social sciences (pp. 175–198). Newbury Park, CA: Sage.

Maki BE. (1997). Gait changes in older adults: Predictors of falls or indicators of fear. J Am Geriatr Soc: 45: 313-320.

Mancini M & Horak FB. (2010). The relevance of clinical balance assessment tools to differentiate balance deficits. Eur J Phys Rehabil Med, 46 (2), 239-248.

Marmot M, Banks J, Blundell R, Lessof C, Nazroo J. (2002). Health, wealth and lifestyles of the older population in England. The 2002 English Longitudinal Study of Ageing. London: IFS.

Maslovskaya O, Smith PWF & Padmadas SS. (2018). A comparison of simple score and latent class approaches: application to HIV knowledge in Chinese and multi-country contexts. International Journal of Social Research Methodology, 21 (2), 243-261. Doi: 10.1080/13645579.2017.1367555.

Mazor KM, Clauser BE, Field T, Yood RA & Gurwitz JH. (2002). A demonstration of the impact of response bias on the results of patient satisfaction surveys. HSR: Health Services Research 37:5: 1403-1417.

Meredith W. (1993). Measurement invariance, factor analysis and factorial invariance. Psychometrika, 58 (4), 525–543. Doi: https://doi.org/10.1007/BF02294825.

Milton K, Vareia AR, Strain T, Cavill N, Foster C & Mutrie N. (2018). A review of global surveillance on the muscle strengthening and balance elements of physical activity recommendations. JFSF, 3 (2), 114-124. Doi: 10.22540/JFSF-03-114.

Mitchell WK, Williams J, Atherton P, Larvin M, Lund J, Nanci M. (2012). Sarcopenia, dynapenia and the impact of advancing age on human skeletal muscle size and strength: a quantitative review. Front Physiol, 12 (3), 260.

Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. Doi:10.1371/journal.pmed1000097.

Mooney SJ, Joshi S, Cerda M, Quinn JW, Beard JR, Kennedy GJ, Benjamin EO, Ompad DC, & Rundle AG. (2015). Patterns of physical activity among older adults in New York city: A latent class approach. Am J Prev Med, 49, 3, e13-e22.

Morris JN & Hardman AE. (1997). Walking to health. Sports Med., 23 (5), 306-332.

Mplus (v7.4) (2016). Muthén, L.K. and Muthén, B.O, Los Angelos, CA. Downloaded on 25th November 2016 from: https://www.statmodel.com/orderonline/

Murphy CB. (2018). What are the pros and cons of stratified random sampling? Downloaded on 19th July 2018 from: https://www.investopedia.com/ask/answers/041615/what-are-advantages-and-disadvantages-stratified-random-sampling.asp

Murphy SL. (2009). Review of physical activity measurement using accelerometers in older adults: considerations for research design and conduct. Prev Med, 48, 108–114.

Murtagh EM, Murphy MH, Murphy NM, Woods C, Nevill AM & Lane A. (2015). Prevalance and correlates of physical activity in community-dwelling older adults in Ireland. PLos ONE, 10(2), e0118293. Doi: 10.1371/journal.pone.0118293.

Muthén LK & Muthén BO. (1998-2017). Mplus User's Guide. Eighth Edition. Los Angeles, CA: Muthén & Muthén Copyright.

Myers AM, Fletcher PC, Myers AH & Sherk W. (1998). Discriminative and evaluative properties of the Activities-specific Balance Confidence (ABC) scale. J. Gerontol.

Myers AM, Powell LE, Maki BE, Holliday PJ, Brawley LR & Sherk W. (1996). Psychological indicators of balance confidence: relationship to actual and perceived abilities. J. Gerontol. A Biol. Sci. Med. Sci, 51.

McCoach DB, Black AC & O'Connell AA. (2007). Errors of inference in structure equation modelling. Psychology in the Schools, 44(5), 461-470. Doi: 10.1002/pits.20238.

McMullan II, McDonough SM, Tully MA, Cupples M, Casson K & Bunting BP. (2018). The association between balance and free-living physical activity in an older healthy community-dwelling adult population: A systematic review and meta-analysis. BMC Public Health, 18, 431.
Doi: https://doi.org/10.1186/s12889-018-5265-4.

National Institute for Health and Care Excellence (NICE). (2013). Falls in older people: assessing risk and prevention. Clinical guidance CG161. Download on 6th June 2018 from:
https://www.nice.org.uk/guidance/cg161/chapter/Introduction.

National Institute for Health and Care Excellence (NICE). (2015). Process and methods guide: Developing NICE guidelines: the manual appendices A-I. Downloaded from: https://www.nice.org.uk/media/default/About/what-we-do/NICE-guidance/NICE-guidelines/developing-NICE-guidelines-the-manual.pdf accessed on 21st April 2016.

Newman, AB, Arnold, AM, Naydeck, BL, Fried, LP, Burke, GL, Enright, P, Gottdiener, J, Hirsch, C, O'Leary, D, Tracy, R, and the Cardiovascular Health Study Research Group. (2003). Cardiovascular health study research group. 'successful ageing': effect of subclinical cardiovascular disease. Arch Intern Med, 163, 2315–2322.

Newtown RA. (2001). Validity of the multi-directional reach test: A practical measure for limits of stability in older adults. The Journals of Gerontology. Series A. 56 (4): m248-m252. Downloaded on 17th Oct 2017 from: https://doi.org/10.1093/gerona/56.4.M248.

Neyer FJ & Lehnart J. (2007). Relationships matter in personality development.: evidence from an 8-year longitudinal study across young adulthood. Journal of Personality, 75(3), 535-568.

Nylund KL, Asparouhov T & Muthén BO. (2007) Deciding on the number of classes in latent class analysis and growth mixture modelling: A monte carlo simulation study. Structural Equation Modelling, 14(4) 535-569, Doi: 10.1080/10705510701575396.

O'Connor, SR, Tully, MA, Ryan, B, Bleakley, CM, Baxter, GD, Bradley, JM & McDonough, SM. (2015). Walking exercise for chronic musculoskeletal pain: systematic review and meta-analysis. Archives of Physical Medicine and Rehabilitation, 96: 724-734.

Office of Disease Prevention and Health Promotion. (2015). https://health.gov/paguidelines/guidelines/appendix1.aspx Accessed on 15th February 2017.

Osuka Y, Yabushita N, Kim M, Seino S, Nemoto M, Jung S, . . . Tanaka K. (2015). Association between habitual light-intensity physical activity and lower-extremity performance: A cross-sectional study of community-dwelling older Japanese adults. Geriatrics & Gerontology International, 15(3), 268-275. Doi: http://dx.doi.org/10.1111/ggi.12268.

Paffenbarger RS Jr, Hyde RT, Wing AL, Hsieh CC. (1986). Physical activity, all-cause mortality and longevity of college alumni. N Engl J Med, 314, 605–613.

Paillard, T, Lafont, C, Costes-Salon, MC, Riviere, D & Dupui, P. (2004). Effects of brisk walking on static and dynamic balance, locomotion, body composition, and aerobic capacity in ageing healthy active men. International Journal of Sports Medicine, 25(7): 539-546.

Paterson DH & Warburton DE. (2010). Physical activity and functional limitations in older adults: a systematic review related to Canada's physical activity guidelines. International Journal of Behavioural Nutrition and Physical Activity, 7, 1-22. Doi:10.1186/1479-5868-7-38.

Perrin PP, Gauchard GC, Perrot C & Jeandel C. (1999). Effects of physical and sporting activities on balance control in elderly people. British Journal of Sports Medicine, 33(2): 121-126.

Physical activity guidelines advisory committee scientific report. (PAGACSR) (2018). Washington, DC: U.S. Department of Health and Human Services, Part G: Research Needs. G2-G44. Downloaded on 2nd May 2018 from: https://health.gov/paguidelines/second-edition/report/.

Piriyaprasarth P, Morris ME, Winter A, Bialocerkowski P. (2008). The reliability of knee joint position testing using electrogoniometry. BMC Musculoskeletal Disorders, 9:6, 1-10. Doi:10.1186/1471-2474-9-6.

Pickard CM, Sullivan PE, Allison GT, Singer KP (2003) Is there a difference in hip joint position sense between young and older groups? J Gerontol A Biol Sci Med Sci 58(7):631–635

Podsialdlo, D & Richardson, S. (1991). The Timed “Up & Go”: A test of basic functional mobility for frail elderly persons. J Am Geriatr Soc, 39: 142-148.

Powell LE & Myers A M. (1995). The Activities-specific Balance Confidence (ABC) Scale. Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 50A (1): M28-34.

Powell KE, Paluch AF & Blair SN. (2011). Physical activity for health: what kind? How much? How intense? On top of what? Annu. Rev. Public Health, 32, 349-365. Doi: 10.1146/annurev-publhealth-031210-101151.

Preston SH, Hill ME, & Drevenstedt GL. (1998). Childhood conditions that predict survival to advanced ages among African–Americans. Social Science and Medicine, 47(9): 1231-1246.

Public Health England (PHE). (2017). Falls: applying all our health. Download on 7th June 2018 from: https://www.gov.uk/government/publications/falls-applying-all-our-health/falls-applying-all-our-health.

Public Health England (PHE). (2018). Muscle and bone strengthening and balance activities for general health benefits in adults and older adults. Download on 24th July 2018 from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/721874/MBSA_evidence_review.pdf.

Ramkissoona I & Cole M (2011). Self-reported hearing difficulty versus audiometric screening in younger and older smokers and non-smokers. J Clin. Med Res: 3(4): 183-190.

Rahal MA, Alonso AC, Andrusaitis FR, Rodrigues TS, Speciali DS, Greve JMDA & Leme LEG. (2015). Analysis of static and dynamic balance in healthy elderly practitioners of Tai Chi Chuan versus ballroom dancing. Clinics (Sao Paulo, Brazil), 70(3), 157-161.

Raykov T & Marcoulides GA. (2000). A first course in structural equation modeling. Mahwah, NJ: Erlbaum. Downloaded on 4th April 2018 from:

file:///C:/Users/Darryl/AppData/Local/Packages/Microsoft.MicrosoftEdge_8wekyb3d8bbwe/TempState/Downloads/9781135600761_googlepreview%20(1).pdf

Refworks (2.0). ProQuest, Mitchigan, US.

Reiner M, Niermann C, Jekauc D & Woll A. (2013). Long term health benefits of physical activity- a systematic review of longitudinal studies. BMC Public Health, 13, 813. http://www.biomedcentral.com/1471-2458/13/813.

Relph N & Herrington L. (2015). Criterion-related validity of knee joint-position-sense measurement using image capture and isokinetic dynamometry. Journal of Sport Rehabilitation, e-pub only http://dx.doi.org/10.1123/jsr.2013-0119.

Revman (2011). Review Manager (RevMan). 5.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration.

Rodgers MM, Pai VM & Convoy RS. (2014). Recent advances in e-wearable sensors for health monitoring. IEE Sensors Journal, 15 (6), 1-8. Doi: 10.1109/JSeN.2014.2357257.

Rose DJ. (2005). In: Jones CJ, Rose DJ editor(s). Physical activity instruction of older adults. Champaign, Illinois: Human Kinetics, 211–227.

Royal College of Physicians. (RCP) (2015). National audit of inpatient falls; audit report. Downloaded on 6th June 2018 from https://www.rcplondon.ac.uk/projects/outputs/naif-audit-report-2015.

Rudyard Kipling (1902). How the camel got its hump. Downloaded from: http://www.kiplingsociety.co.uk/poems_hump.htm. Accessed on 2nd May 2018.

Sabia S, Singh-Manoux A, Hagger-Johnson G, Cambois E, Brunner E & Kivimaki M. (2012). Influence of individual and combined healthy behaviours on successful ageing. CMAJ, Canadian Medical Association, 184 (18), 1985-1992.

Saelens BE, Sallis JF, Frank LD, Cain KL, Conway TL, Chapman JE, Slymen DJ & Kerr J. (2012). Neighborhood environmental and psychosocial correlates of adults' physical activity. Med Sci Sports Exerc, 44, 637–646.

Sakurai R, Fujiwara Y, Ishihara M, Higuchi T, Uchida H & Imanaka K. (2013). Age-related self-overestimation of step-over ability in healthy older adults and its relationship to fall risk. BMC Geriatrics 13,44. http://www.biomedcentral.com/1471-2318/13/44.

Salter K, Jutai JW, Teasell R, Foley NC, Bitensky J, Bayley M. (2005). Issues for selection of outcome measures in stroke rehabilitation: ICF activity. Disabil. Rehabil. 27, 315–340.

Sanderson S, Tatt ID & Higgins JPT. (2007). Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. International Journal of Epidemiology. 36: 666–676. Doi:10.1093/ije/dym018.

Santos Mendes FA, Caromano FA, Ide MR, Schujmann DS, Almeida MHM & Carvalho EV. (2011). General versus walking exercises on the static and dynamic balance of healthy elderly persons. Ter Man, 9 (46), 167-174.

Saris WE & Stronkhorst H. (1984). Casual modelling in non-experimental research: an introduction to the LISREL approach. Sociometric Research Foundation, Amsterdam.

Savela SL, Koistinen P, Stenholm S, Tilvis RS, Strandberg AY, Pitkala KH, Salomaa VV & Strandberg TE. (2013). Leisure-time physical activity in mid-life is related to old age frailty. Journal of Gerontology: Medical Sciences, 68 (J1), 1433-1438.

Schafer JL & Graham JW. (2002). Missing data: our view of the state of the art. Psychological Methods., 7(2), 147-177.

Schmidt SEC, Tittlbach, Bos K & Well A. (2017). Different types of physical activity and fitness and health in adults: An 18-year longitudinal study. BioMed Research International, Article ID 1785217, 1-10 Downloaded from: https://doi.org/10.1155/2017/1785217. Accessed on 26th April 2018.

Schoppen T, Boonstra A, Groothoff JW, de Vries J, Goeken LNH & Eisma WH. (1999). The timed "Up and Go" test: reliability and validity in persons with unilateral lower limb amputation. Arch Phys Med Rehabil, 80, 825-828.

Schwartz, G. (1978). Estimating the dimension of a model. The Annals of Statistics, 6, 461–464.

Schutzer KA & Graves S. (2004). Barriers and motivations to exercise in older adults. Preventive Medicine, 39,1056–1061.

Shephard RJ. (2003). Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med, 37(3).197–206.

Sherrington C, Michaleff ZA, Fairhall N, et al. (2017). Exercise to prevent falls in older adults: an updated systematic review and meta-analysis. Br J Sports Med 7; 51:1750-1758.

Shumway-Cook A & Woollacott MH. (2001). Motor Control: theory and practical applications. Philadelphia: Lippincott, Williams & Wilkins.

Shumway-Cook A & Woollcott MH. (2007). Motor control: translating research into clinical practice. 3rd Edition Philadelphia: Lippincott Williams & Wilkins.

Sibley KM, Beauchamp MK, Van Ootghem K, Straus SE & Jaglal SB. (2015). Using the systems framework for postural control to analyse the components of balance evaluated in standardized balance measures: a scoping review. Archives of Physical Medicine and Rehabilitation, 96, 122-132.

Silverwood RJ, Nitsch D, Pierce M, Kuh D, & Mishra GD. (2011). Characterizing longitudinal patterns of physical activity in mid-adulthood using latent class analysis: Results from a prospective cohort study. Am J Epidemiol., 174,12, 1406–1415. Doi: 10.1093/aje/kwr266.

Skelton DA & Mavroeidl A. (2018). How do muscle and bone strengthening and balance activities (MBSBA) vary across the life course, and are there particular ages where MBSBA are most important? JFSF, 3 (2), 74-84. Doi: 10.22540/JFSF-03-074.

Smith AK, Ayanian JZ, Covinsky KE, Landon BE, McCarthy EP, Wee CC & Steinman MA. (2011). Conducting high value secondary dataset analysis: an introductory guide and resources. J Gen Intern Med, 26(8), 920-929.

Sofi F, Valecchi, Bacci D, Abbate R, Gensim GF, Casini A & Macchi C. (2011). Physical activity and risk of cognitive decline: a meta-analysis of prospective studies. Journal of Internal Medicine, 269, 107-117. Doi:10.1111/j.1365-2796.2010.02281.x.

Sorbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. Br. J. math. Statist. Psychol., 27(2), 229-239.

Stephenson MT, Holbert RL & Zimmerman RS. (2006). On the use of structural equation modelling in health communications research. Health Communication, 20(2), 159-167.

Steptoe A, Breeze E, Banks J, Nazroo J. (2013). Cohort profile: the English Longitudinal Study of Ageing. Int J Epidemiol. 42 (6), 1640-1648. Doi: 10.1093/ije/dys168 PMID: 23143611.

Stoel RD, van Den Wittenboer G, & Hox J. (2003). Analysing longitudinal data using multilevel regression and latent growth curve analysis. Metodologia de las ciencias del comportamiento. 5: 1-21.

Strain T, Fitzsimmons C, Kelly P & Mutrie N. (2016). The forgotten guidelines cross-sectional analysis of participation in muscle strengthening and balance & coordination activities by adults and older adults in Scotland. BMC Public Health, 16, 1108. Doi: 10.1186/s12889-016-3774-6.

Stubbs B, Brefka S, & Denkinger MD. (2015). What works to prevent falls in community dwelling older adults? Umbrella review of meta-analyses of randomized controlled trials. Physical Therapy, 95 (8): Downloaded from: http://ptjournal.apta.org/content/early/2015/05/14/ptj.20140461.

Studenski, S. (2009). Bradypedia: is gait speed ready for clinical use? J Nutr Health Aging, 13(10), 878–880.

Suri P, Kiely DK, Leveille SG, Frontera WR, & Bean JF. (2011). Increased trunk extension endurance is associated with meaningful improvement in balance among older adults with mobility problems. Arch. Phys. Med. Rehabil., 92, 1038–1043.

Tak E, Kuiper R, Chorus A, & Hopman-Rock M. (2013). Prevention of onset and progression of basic ADL disability by physical activity in community dwelling older adults: A meta-analysis. Ageing Research Reviews 12:329-228.

Tangen GG, Engedal K, Bergland A, Moger TA, & Mengshoel AM. (2014). Relationships Between Balance and Cognition in Patients With Subjective Cognitive Impairment, Mild Cognitive Impairment, and Alzheimer Disease, Physical Therapy, Volume 94, Issue 8, 1123–1134, Doi: https://doi.org/10.2522/ptj.20130298

Theou O, Stathokostas L, Roland KP, Jakobi JM, Patterson C, Vandervoort AA & Jones GR. (2011). The effectiveness of exercise interventions for the management of frailty: a systematic review. Journal of Ageing Research. Vol 2011, Article ID 569194, 19 pages. Doi: 10.406/2011/569194.

Thomas JC, Odonkor C, Griffith L, Holt N, Percac-Lima S, Leveille S, Ni P, Latham NK, Jette AM, & Bean JF. (2014). Reconceptualizing balance: attributes associated with balance performance. Experimental Gerontology, 57, 218-223.

Tierney MC, Moineddin R, Morra A, Manson, J, & Blake. J. (2010). Intensity of recreational physical activity throughout life and later life cognitive functioning in women. J Alzheimers Dis., 22(4), 1331-1338.

TILDA Derived Variables Codebook wave 1 v1.5 http://www.tilda.ie. Downloaded on 26th November 2016.

TILDA Derived Variables Codebook wave 2 v2.2 http://www.tilda.ie. Downloaded on 26th November 2016.

TILDA Release Guide (2016) version 2.0. http://www.tilda.ie. Downloaded on 26th November 2016.

Tinetti ME, Inouye SK, Gill TM & Doucette JT. (1995). Shared risk factors for falls, incontinence and functional dependence: unifying the approach to geriatric syndromes. JAMA, 273: 1348–1353.

Tomarken AJ & Waller NG. (2005). Structural equation modelling: strengths, limitations, and misconceptions. Annu. Rev. Clin. Psychol. 1, 31-65. Doi: 10.1146/annurev.clinpsy.1.102803.144239.

Toonstra J & Mattacola CG. (2012). Test–retest reliability and validity of isometric knee-flexion and extension measurement using 3 methods of assessing muscle strength. Journal of Sport Rehabilitation, 1-5.

Toshiaki T, Kenji I, Haruyasu Y, Jun T, Masanor, N, Yoshinori D & Hiroshi Y. (2006). Modification of the functional reach test: Analysis of lateral and anterior functional reach in community-dwelling older people. Archives of Gerontology and Geriatrics, 42, 167-173.

Tsang WWN & Hui-Chan CWY. (2004). Effects of exercise on joint sense and balance in elderly men: Tai Chi versus golf. Medicine & Science in Sports & Exercise, 36(4), 658-667 10p.

Tsang WWN & Hui-Chan CWY. (2005). Comparison of muscle torque, balance, and confidence in older Tai Chi and healthy adults. Medicine and Science in Sports and Exercise, 37(2), 280-289.

Tsang WWN & Hui-Chan CWY. (2006). Standing balance after vestibular stimulation in Tai Chi-practicing and nonpracticing healthy older adults. Archives of Physical Medicine & Rehabilitation, 87(4), 546-553 8p.

Tsang WWN & Hui-Chan CWY. (2010). Static and dynamic control in older golfers. Journal of Aging and Physical Activity, 18, 1-13.

Tsang WWN, Wong VS, Fu SN & Hui-Chan CWY. (2004). Tai Chi Improves Standing Balance Control under Reduced or Conflicting Sensory Conditions. Archives of Physical Medicine and Rehabilitation, 85(1), 129-137.

Tse ACY, Wong TWL & Lee PH. (2015). Effect of low-intensity exercise on physical and cognitive health in older adults: a systematic review. Sports Medicine, 1 (37), 1-13. Doi: 10.1186/s40798-015-0034-8.

Tu YK. (2009). Commentary: is structural equation modelling a step forward for epidemiologists. Int J Epidemiol, 38, 1-3.

Tyson SF & Connell LA. (2009). How to measure balance in clinical practice: a systematic review of the psychometrics and clinical utility of measures of balance activity for neurological conditions. Clin Rehabil, 23, 824-840.

Umpierre D, Kramer CK, Leita CB, et al. (2011). Physical activity advice only or structured exercise training and association with HbA1c levels in type 2 diabetes. JAMA, 305, 1790-1799.

Vaillant GE & Mukamal K. (2001). Successful ageing. Am J Psychiatry, 158, 839–847.

Valete-Rosalino CM & Rozenfeld S. (2005). Auditory screening in the elderly: comparison between self-report and audiometry. Rev Bras Otorrinolaringol, 71(2), 193-200.

Van de Schoot R, Lugtig PJ & Hox JJ. (2012). A checklist for testing measurement invariance. European Journal of Developmental Psychology, 9 (4), 486-492.

Van Tulder MW, Esmail R, Bombardier C, Koes BW. (2000). Back schools for non-specific low back pain. Cochrane Database Syst Rev. (2), CD000261.

Varma VR, Tan EJ, Wang T, Xue Q, Fried LP, Seplaki CL, Carlson MC. (2014). Low-intensity walking activity is associated with better health. Journal of Applied Gerontology: The Official Journal of the Southern Gerontological Society, 33(7), 870-887. Doi: http://dx.doi.org/10.1177/0733464813512896.

Vieira ER, Palmer RC & Chaves PHM. (2016). Prevention of falls in older people living in the community. BMJ, 353, i1419. Doi:10.1136/bmji1419.

Visser JE, Carpenter MG, van der KH, & Bloem BR. (2008). The clinical utility of posturography. Clin.Neurophysiol, 119, 2424–2436.

Wayne PM, Gow BJ, Costa MD, Peng CK, Lipsitz LA, Hausdorff JM, Davis RB, Walsh JN, Lough M, Novak V, Yeh GY, Ahn AC, Macklin EA & Manor B. (2014). Complexity-based measures inform effects of Tai Chi training on standing postural control: cross-sectional and randomised trial studies. PLoS ONE, 9(12), e114731. Doi: 10:1371/journal.pone.0114731.

Welk GJ. (2002). Physical activity assessments for health-related research. Human Kinetica. Human Kinetics Publishers, Champaign. Il.

Wells GA, Shea B, O'Connell D, Robertson J, Peterson J, Welch V, Logos M & Tugwell, P. (2010) The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-analyses, Ottawa Hospital Research Institute, Ottawa, Canada. Downloaded on 16th March 2017 from: http://www.evidencebasedpublichealth.de/download/Newcastle_Ottowa_Scale_Pope_Bruce.pdf.

Wen CP, Wai JP, Tsai MK et al. (2011). Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. Lancet, 378, 1244–1253.

Whitney SL, Wrisley DM, Marachetti GF, Gee MA, Redfern MS & Furman J. (2005). Clinical measurements of sit-to-stand performance in people with balance disorders: validity of data for the five-times-sit-to-stand test. Phys Ther 85(10): 1034-1045.

Winter DA. (1995). ABC: Anatomy, biomechanics and control of balance during standing and walking. Waterloo, Ont: Waterloo Biomechanics, 1995.

Wirtz A & Nachreiner F. (2010). The effects of extended working hours on health and social well-being-a comprehensive analysis of four independent samples. Chronobiol Int., 20 (5), 1124-1134.

Wong AMK, Lin Y, Chou S, Tang F & Wong P. (2001). Coordination exercise and postural stability in elderly people: effect of Tai Chi Chuan. Archives of Physical Medicine & Rehabilitation, 82(5), 608-612 5p.

Wong AMK, Chou S, Huang S, Lan C, Chen H, Hong W, Chen CPC & Pei Y. (2011). Does different exercise have the same effect of health promotion for the elderly? Comparison of training-specific effect of Tai Chi and swimming on motor control. Archives of Gerontology and Geriatrics, 53(2), e133-e137.

Worldwide Health Organisation. (WHO). (2002). Active ageing: A policy framework. (WHO/NMH/NPH/02.8), Geneva: WHO non-communicable diseases and mental health clusters. Downloaded on 25th April 2018 from: http://apps.who.int/iris/bitstream/handle/10665/67215/WHO_NMH_NPH_02.8.pdf;jsessionid=59BADA5D0A8816E8AAB4B96F52188A79?sequence=1.

Worldwide Health Organisation. (WHO). (2007). WHO Global report on falls prevention in older age. Downloaded on 27th September 2015 from: http://www.who.int/ageing/publications/Falls_prevention7March.pdf.

Worldwide Health Organisation. (2010). Global recommendations on physical activity for health. Downloaded on 20 June 2017 from:

http://www.who.int/dietphysicalactivity/factsheet_recommendations/en/.

Worldwide Health Organization (WHO). (2011). Global recommendations on physical activity for health. 65 years and above. Downloaded on 2nd May 2017 from: http://who.int/dietphysicalactivity/physical-activity-recommendations-65years.pdf?ua=1.

Worldwide Health Organisation (WHO). (2013) Global Physical Activity Questionnaire (GPAQ) analysis guide. Downloaded on 6th April 2016 from: http://www.who.int/chp/steps/resources/GPAQ_Analysis_Guide.pdf?ua=1.

Worldwide Health Organisation (WHO). (2015). World Report on ageing and health. Downloaded on 2nd November 2015 from: http://apps.who.int/iris/bitstream/10665/186463/1/9789240694811-eng.pdf?ua=.

Worldwide Health Organisation (WHO). (2018). Falls. Downloaded on 6th June 2018 from: http://www.int.news-room/fact-sheets/detail/falls.

Worldwide Health Organisation (WHO). (2018). Global strategy on diet, physical activity, and health: Intensity of physical activity. Downloaded on 7th June 2018 from: https://www.who.int/dietphysicalactivity/physical_activity_intensity/en/.

Xu H & Tracey TJG. (2017). Use of Multi-Group Confirmatory Factor Analysis in Examining Measurement Invariance in Counselling Psychology Research. The European Journal of Counselling Psychology, 6(1), 75-82. Doi: http://dx.doi.org/10.5964/ejcop.v5i2.120.

Yang Y, Verkuilen JV, Rosengren KS, Grubisich SA, Reed MR & Hsiao-Wecksler ET. (2007). Effect of combined Taiji and Qigong training on balance mechanisms: a randomized controlled trial of older adults. Medical Science Monitor 13(8), R339-348.

Young J, Angevaren M, Rusted J & Tabet N. (2015). Aerobic exercise to improve cognitive function in older people without known cognitive impairment. Cochrane Database of Systematic Reviews, Issue 4. Art. No. CD005381. Doi: 10.1002/14651858.CD005381.pub4.

Yuan KH & Bentler PM. (2000). Three likelihood-based methods for mean and covariance structure analysis with non-normal missing data. Sociological Methodology, 165-200.

Zhang C, Mao D, Riskowski JL & Song Q. (2011). Strategies of stepping over obstacles: The effects of long-term exercise in older adults. Gait & Posture, 34(2), 191-196.

Zhu W, Wadley VG, Howard VJ, Hutto B, Blair SN & Hooker SP. (2016). Objectively measured physical activity and cognitive function in older adults. Medicine & Science in Sports & Exercise., 47-53. Doi:10.1249/MSS.0000000000001079. Downloaded from: http://www.acsm-msse.org.

Zimdars A, Nazroo J & Gjonça E. (2011). The circumstances of older people in England with self-reported visual impairment: A secondary analysis of the English Longitudinal Study of Ageing (ELSA). The British Journal of Visual Impairment 30(1), 22 –30. Doi: 10.1177/0264619611427374 jvi.sagepub.com.

Appendices

Appendix I
Table summarising changes made between protocol and systematic review study

Location of change	Details of change	Rationale
Types of studies	Observational and RCTs that included a comparison group regarding physical activity were included in the types of studies.	To ensure that the differences in active and less active groups could be assessed.
Types of intervention	Additional details regarding exercise interventions was included, where structured and planned exercise that took place in a researcher environment or a healthcare facility was excluded.	To remove any exercise that was not carried out for the purposes of leisure activity i.e. sports field/club.
Search strategy	Inclusion of searches of longitudinal studies such as TILDA, ELSA, and NICOLA and the National Institute for Health Research library.	To ensure any appropriate studies, publications or reports that had not been identified in the main search were included in the review.
Risk of bias assessment for observational studies	Use of the Newcastle Ottawa Scale rather than Down & Black.	Downs and Black were trialled using a couple of the included papers, but the questions were found not relevant for cross-sectional studies.

Appendix II
Medline Search

(The following search with adjustments was implemented in the additional databases searched).

1. musculoskeletal physiological phenomena/ or postural balance/

2. Accidental Falls/

3. 1 or 2

4. human activities/ or "activities of daily living"/ or exercise/ or leisure activities/ or travel/ or work/

5. movement/ or gait/ or running/ or swimming/ or walking/ or physical endurance/ or physical fitness/

6. 4 or 5

7. 3 and 6

8. (healthy not (amput* or arthriti* or osteoporos* or "musc* dis*" or "nerv* system dis*" or "neur* dis*" or Alzheimer* or Parkinson* or dementia* or "multiple sclerosis*" or "somatosensory* dis*" or "hear* dis*" or "vis* dis*" or "history of fall*" or "history of fracture*")).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]

9. 7 and 8

10. limit 9 to (english language and ("middle age (45 to 64 years)" or "middle aged (45 plus years)" or "all aged (65 and over)" or "aged (80 and over)"))

Appendix III
Cochrane risk of bias form used for intervention studies

Risk of bias: RCTS
Paper: Reviewer: Date:

Risk of bias	Risk (Please circle)	Support for judgement
Random sequence generation (selection bias)	LOW UNCLEAR HIGH	
Allocation concealment (selection bias)	LOW UNCLEAR HIGH	
Blinding of participants and personnel (performance bias)	LOW UNCLEAR HIGH	
Blinding of outcome assessment (detection bias)	LOW UNCLEAR HIGH	
Incomplete outcome data (attrition bias)	LOW UNCLEAR HIGH	
Selective reporting (reporting bias)	LOW UNCLEAR HIGH	
Other bias	LOW UNCLEAR HIGH	

Appendix IV
Newcastle Ottawa (NOS) Risk of Bias Form used for Observational Studies

Study ID (e.g. Smith, 2000)		Citation	
Date of assessment		Name of assessor	
Overall Rating			

Area	Selection	Rationale
Selection: (Maximum 5 stars)		
1) Representativeness of the sample: a) Truly representative of the average in the target population. * (all subjects or random sampling) b) Somewhat representative of the average in the target population. * (non-random sampling) c) Selected group of users d) No description of the sampling strategy		
2) Sample size: Justified and satisfactory. * Not justified		
3) Non-respondents: a) Comparability between respondents & non-respondents characteristics is established, & the response rate is satisfactory. * b) The response rate is unsatisfactory, or the comparability between respondents and non-respondents is unsatisfactory c) No description of the response rate or the characteristics of the responders and the non-responders.		
4) Ascertainment of the exposure (risk factor): a) Validated measurement tool. ** b) Non-validated measurement tool, but the tool is available or described. * c) No description of the measurement tool.		
Comparability (maximum 2 stars)		
The subjects in different outcome groups are comparable, based on the study design or analysis. Confounding factors are controlled. a) The study controls for the most important factor (select one). * b) The study control for any additional factor. *		
Outcome (maximum 3 stars)		
1) Assessment of the outcome: a) objective validated assessment. ** b) objective non-validated assessment. ** c) Self-report. * d) No description.		
2) Statistical test: a) The statistical test used to analyze the data is clearly described and appropriate, and the measurement of the association is presented, including confidence intervals and the probability level (p value). * b) The statistical test is not appropriate, not described or incomplete.		

Original form Wells et al., 2010 and adapted for cross-sectional studies by Herzog et al., 2013.

Appendix V
Data extraction form

Study ID	
Date form completed *(dd/mm/yy)*	
Name/ID of person extracting data	
Reference citation	

Eligibility

Confirm eligibility for review	Yes or No
Reason for exclusion	

Characteristics of included study

Methods	Description	Location in text
Aim of study		
Study Design		
Duration of study		

Participants

Total no. participants *(Male/female)*		
Mean age +/- SD (range)		
Inclusion criteria		
Exclusion criteria		
Co-morbidities/ general health		
Socio-demographics (education level)		
Setting (community or institution)		
Method of recruitment of participants		

Other Information

Study funding sources	
Possible conflicts of interest *(for authors)*	

Appendix V (continued)
Data extraction form

Type of Physical Activity (PA):

	Description			Location
Group name (if specified)	G1	G2	G3	
Description of PA measure				
Results of PA measure				
Delivery Setting				
Providers				
Co-interventions (if applicable)				
No. in group (specify if randomised & how)				
Gender (F/M)				
No. of dropouts/withdrawal (reasons)				
Mean age (SD) (Range)				
Participants excluded from study (before or after randomisation)?				
Sample size calculation				
Unit of randomisation				
Baseline differences				
Adverse events				
Notes:				

Balance measures

Balance measure		
Description		
Time points measured (specify from start/end of intervention)		

Appendix V (continued)
Data extraction form

	Description	Location
Person measuring		
Unit of measurement (if relevant)		
Measurement tool *(indicate if high or low score is good)*		
Is outcome/tool validated?		

Results

Any other results reported *(e.g. mean difference, CI, P value)*		
Imputation of missing data (e.g. assumptions made for ITT analysis)		
Statistical methods used and appropriateness of these		
Key conclusions of study authors		
References to other relevant studies		
Study strengths		
Study limitations		
Recommendations made by study		

(adapted using criteria from Higgins & Green, 2011)

Appendix VI
Formulae for combining groups

(Higgins and Green, 2011, Part 2 General Methods for Cochrane Reviews, Chapter 7)

	Group 1 (e.g. males)	Group 2 (e.g. females)	Combined groups
Sample size	N_1	N_2	$N_1 + N_2$
Mean	M_1	M_2	$\frac{N_1 M_1 + N_2 M_2}{N_1 + N_2}$
SD	SD_1	SD_2	$\sqrt{\frac{(N_1 - 1) SD_1^2 + (N_2 - 1) SD_2^2 + \frac{N_1 N_2}{N_1 + N_2}(M_1^2 + M_2^2 - 2M_1 M_2)}{N_1 + N_2 - 1}}$

Appendix VII
Detailed description of each balance measures included in all studies (observational and intervention studies)

Primary outcome measures

Neuromuscular System: measures of gait

Walking speed

Walking speed, is a core indicator of health and function in ageing and disease (Studenski et al., 2009) and is a good predictor for major health related outcomes such as falls (Abeelan van Kan et al., 2009; Graham et al., 2008).

Preferred walking speed, the time to walk a pre-determined distance, was used in 4 observational studies (284 participants) (Aoyagi et al., 2009; Dewhurst et al., 2014; Gaudagnin et al., 2015; and Zhang et al., 2011), and 2 RCTs (81 participants) (Paillard et al., 2004; Wayne et al., 2015).

Maximal walking speed, the maximum time taken to walk a pre-determined distance, was used in 2 observational studies (230 participants) (Aoyagi et al., 2009; Dewhurst et al., 2014).

The distance of walk varied between studies between 5 – 75 metres and was undefined in 1 observational study (Zhang et al., 2011) and 1 RCT (Paillard et al., 2004). This was expressed as velocity (m/s). A higher value of velocity indicates better balance performance.

Neuromuscular system: measures of strength

Handgrip test

Handgrip strength is a measure for overall body muscle strength and a predictor of disability and mobility limitations (Rananen et al., 1999; Shinkaiet et al., 2000). Handgrip test measures the maximum isometric strength of the hand and forearm muscles using a dynamometer. 2 observational studies (210 participants) measure handgrip strength (Aoyagi et al., 2009; Tsang et al., 2004) using 2 different dynamometers (Smedley: Aoyagi et al., 2009; Jamar: Tsang et al., 2004) but research suggests that regardless of type of dynamometer, the results are reliable and comparable (Schmidt et al., 2002). Both studies use different units of measurement: peak force (N) and weight (Kg) and so N were converted to Kg. A higher value indicates better balance performance.

Appendix VII (continued)
Detailed description of each balance measures included in all studies (observational and intervention studies)

Isometric Knee Extension

The isometric knee extension test measures lower limb muscle strength and is measured in 4 observational studies (Aoyagi et al., 2009; Brooke-Wavell & Cooling, 2008; Gauchard et al., 1999; Tsang & Hui-Can, 2005) (218 participants) using a dynamometer in 3 studies (Aoyagi et al., 2009: Tas; Gauchard et al., 1999: Biodex Corp; Tsang & Hui-Chan, 2005: Cybex norm) and in 1 study using a purpose-built scat and force meter (Brooke-Wavell & Cooling, 2008). The test results were expressed in force (N), velocity (m/Kg; NM/s), and peak torque to body weight ratio (NM/Kg). The knee joint was held at a 90-degree angle where higher values indicated better balance performance.

Ultrasound test

2 observational studies use ultrasound to measure bone density (Brooke-Wavell & Cooling, 2008; Fong et al., 2014). There is variation in terms of the tests where Brooke-Wavell & Cooling (2008) use a Broadband Ultrasonic Attenuation of the calcaneus (ankle bone) using an Osteometer (dB/MHz). Fong et al. (2014) measures bone strength of the distal radius of the dominant arm using a sonometer measuring velocity (SOS). High values indicate better balance performance.

<u>Neuromuscular system: measures of functionality</u>

Chair stand test

The chair stand test, a measurement to assess functional lower extremity strength in older adults, is measured in 1 observational study (Hakim et al., 2004) (65 participants). Results were expressed in time (s) to complete 5 sit and stands, number of full sit and stands (n) in 30 seconds, and time take for weight transfer (s) moving from sitting to standing position. Higher values indicate better balance performance.

Single leg jump test

The single leg jump test is a measure of strength and balance control and involves jumping on one leg high enough to leave the floor from a starting position of a single-leg stance with eyes open for as long as possible up to a maximum of 30 seconds. 1 observational study used this test (Gyllensten et al., 2010) (44 participants) and a high value indicates better balance performance.

Appendix VII (continued)
Detailed description of each balance measures included in all studies (observational and intervention studies)

Timed Up & Go (TUG)

This test measures the time to stand, walk a pre-determined distance, turn, and return to a sitting position in seconds (Podsialdo & Richardson, 1991). It was used in 4 observational studies (286 participants) (Brooke-Wavell & Cooling, 2008; Dewhurst et al., 2014; Hakim et al., 2004; Wayne et al., 2014) and 1 RCT study (Wayne et al., 2014). 2 studies used 3 meters walk (Brooke-Wavell & Cooling, 2008; Hakim et al., 2004); 1 study used a 2.44m walk (Dewhurst et al., 2014) and 1 RCT study did not specify the distance (Wayne et al., 2014). Lower values on these tests indicate a better balance performance.

Fullerton Advanced Balance scale (FAB)

The FAB (Rose et al., 2006) measures physical performance across 10 dynamic standing activities. 2 items are considered low functioning, 6 are moderate, and 2 are high level physical activities. 1 observational study (Hakim et al., 2010) (41 participants) uses the FAB and a high score indicates better balance performance.

Timed Floor Transfer (TFT)

TFT (Murphy et al., 2003) measures the time required to transfer from a standing position to the floor and then to return to a standing position. 1 observational study (Hakim et al., 2010) (41 participants) uses the TFT and a low score indicates better balance performance.

Body Awareness Scale – Health (BAS-H)

BAS-H measures quality of movement, functional ability and balance control and is a 25-item scale. 1 observational study (Gyllensten et al., 2010) (44 participants) uses BAS-H and a low score indicates better balance performance.

Single leg stance (SLS)

Single legged stance is the ability to balance on one leg measured as the time before placing the opposite leg on the ground and was measured across 4 observational studies (Hakim et al., 2010; Tsang & Hui-Chan, 2010; Wayne et al., 2014; Zhang et al., 2011) (181 participants). The condition of eyes open was used across all studies. 2 studies used 30 secs timeframe (Hakim et al., 2010; Tsang et al., 2010), and 1 did not specify time but measured the time spent standing on one foot as the other foot crossed an obstacle (Zhang et al., 2011). Higher values indicate better balance performance.

Appendix VII (continued)

Detailed description of each balance measures included in all studies (observational and intervention studies)

Berg Balance Scale

The Berg Balance Scale is a 56-point scale comprising of 14 items of activities of daily living. Each item is scored from between 0 to 4 (Berg, 1992). This was used in 1 observational study (Fong et al., 2014) and 1 RCT (Yang et al., 2007) (49 participants). Higher values indicate better balance performance.

Neuromuscular system: measures of stability

Functional reach

Functional reach test measures the distance (m, cm, inches) an individual can reach while maintaining a fixed base of support in a standing or seated position (Duncan, 1990) and this was measured in 4 observational studies (Aoyagi et al., 2009; Gao et al., 2011; Hakim et al., 2004; Hakim et al., 2010) 299 participants). Forward reach only was explored in 2 studies (Aoyagi et al., 2009; Gao et al., 2011), and multidirectional reach (forward, backward, left and right) was explored in 2 studies (Hakim et al., 2004; Hakim et al., 2010). Distance (m) was measured in all studies excluding 1 observational study that normalised distance with body height (%) (Gao et al., 2011). Higher results indicate better balance performance.

Neuromuscular system: measures of flexibility

Range of Motion (ROM):

Range of motion is measured across a variation of joints in 2 observational studies (Brooke-Wavell & Cooling, 2008; Dewhurst et al., 2014) (134 participants): Back and hamstring where the higher values of distance between extended fingers and tip of toes (cm) indicate better balance performance; left and right shoulder where higher values of distance between the extended fingers of the 2 hands (cm) indicate poor balance performance; and shoulder and ankle where higher range of motion (degrees) indicate better balance performance.

Cognitive system

Mini Mental State Exam (MMSE)

The MMSE (Folstein et al.,1975), is a 30-item scale measuring attention, concentration, memory, language, Visio-constructional skills, calculations, and orientation and provides a summary score of a maximum of 30 for each

Appendix VII (continued)
Detailed description of each balance measures included in all studies (observational and intervention studies)

participant. 5 observational studies (229 participants) (Gyllensten et al., 2010; Lu et al., 2013; Tsang et al., 2004; Tsang & Hui-Chan, 2004; Wayne et al., 2014) and 2 RCTs (120 participants) (Wayne et al., 2014; Wayne et al., 2015) use the MMSE. 1 study reported the results in median only (Tsang & Hui-Chan, 2004), and 1 study (Lu et al., 2013) did not report the version or number of items used. A higher value indicates a better balance performance.

Activities Specific Balance Confidence scale (ABC)

The ABC is a subjective measure of confidence in performing various ambulatory activities without falling or experiencing a sense of unsteadiness with a maximum score of 16 for each participant. 4 observational (Fong et al., 2014; Gao et al., 2011; Hakim et al., 2004; Tsang & Hui-Chan, 2005) all use the ABC, of which 2 used a modified version (Fong et al., 2014; Gao et al., 2011). A higher value indicates better balance performance.

Reaction time

The reaction time taken to respond to a light stimulus given at random spacing was tested in 3 observational studies (Brooke-Wavell & Cooling, 2008; Lu et al., 2013; Wong et al., 2011) (198 participants) where all reported time (s) and 1 reported velocity (ms) (Wong et al., 2011). A low value indicates better balance performance.

<u>Sensory system: Proprioception</u>

Knee joint repositioning

General methods for testing joint proprioception include limb segment repositioning using the knee joint repositioning which can be tested in either passive or active mode. 2 observational studies (Fong & Ng, 2006; Tsang & Hui-Chan, 2004) (58 participants) use knee joint repositioning, where 1 uses a passive form in a non-weight bearing condition (Tsang & Hui-Chan, 2004) measuring absolute angle error with a dynamometer and 1 uses an active knee joint repositioning weight bearing test (Fong & Ng, 2006) measuring absolute angle error using a electrogoniometer. Low values indicate good balance performance.

Appendix VII (continued)
Detailed description of each balance measures included in all studies (observational and intervention studies)

<u>Sensory system: vestibular function</u>

Vestibular tests

1 observational study measured vestibular function using caloric and rotational tests where a high score indicates good balance (degrees/s) (Gauchard et al., 2001) (25 participants).

<u>Other</u>

Falls

The number of falls in the last 6 months was recorded for 1 observational study (Brooke-Wavell & Cooling, 2008) (74 participants) where a high score indicates poor balance performance.

Secondary outcome measures

Sensory Organisational Test (SOT)

The SOT measures ability to use visual, vestibular and somatosensory inputs and to suppress sensory information that is inappropriate. The participant stands steady during 3 trials involving 6 sensory conditions: eyes open standing on a fixed surface using a fixed visual surround; eyes closed standing on a fixed surface; eyes open standing on a fixed surface using a sway referenced visual surround; eyes open standing on an uneven platform using a fixed visual surround; eyes closed standing on an uneven surface; and eyes open standing on an uneven surface with a sway referenced visual surround. The test is performed using a computer programme and force platform. 3 observational studies (Buatois et al., 2007; Gao et al., 2011; Tsang et al., 2004) (139 participants) and 1 RCT (Yang et al., 2007) (49 participants) used the SOT. SOT outcome measures include equilibrium and composite scores measuring average centre of gravity for each condition and as a weighted average of individual scores respectively (Buatois et al., 2007), and sensory analysis ratios for each condition which identify impairments of the sensory system (Gao et al., 2011; Tsang et al., 2004). Higher values indicate better balance performance.

Appendix VII (continued)
Detailed description of each balance measures included in all studies (observational and intervention studies)

Force platform and sway indicators: static and dynamic balance

Force platforms allow the measurement of movement of the centre of pressure or limits of stability under different conditions (eyes closed; eyes open) using distance (m), speed (cm/s), area (cm2), and angle (degrees) as total body sway, anterior posterior sway, and mediolateral sway. Force platforms for the measurement of sway for static or dynamic balance was used in 16 observational (Aoyagi et al., 2009; Brooke-Wavell & Cooling, 2008; Dewhurst et al., 2014; Gao et al., 2011; Gauchard et al., 2001; Gauchard et al., 2003; Perrin et al., 1999; Gyllensten et al., 2010; Lu et al., 2013; Rahal et al., 2015; Tsang & Hui-Chan, 2005; Tsang & Hui-Chan, 2006; Tsang & Hui-Chan, 2010; Wayne et al., 2014; Wong et al., 2001; Wong et al., 2011), and 3 RCTs (Felipe et al., 2011; Paillard et al., 2004; Wayne et al., 2014). Low values under static and dynamic conditions measuring sway indicate better balance performance and high values on maximum excursion of loss of stability indicate better balance performance.

Tilt board

The ability to maintain balance whilst standing on a tilt board measured in time to loss of balance (s) was used in 1 observational study (Fong & Ng, 2006). Higher values indicate better balance performance.

Appendices VIII
A comparative analysis of the measures across ageing studies in the UK and Ireland

Relevant outcome measure	NICOLA	ELSA	TILDA
Time points	One	Six	Two
Age/DOB	CAPI	CAPI	CAPI
Gender	CAPI	CAPI	CAPI
Education	CAPI	CAPI	CAPI
Physical Activity	CAPI (7 & 14 days-self-reported and information relating to PA using IPAQ and RPAQ.	CAPI (7 days self-reported PA using a version of IPAQ).	CAPI (7 days self-reported PA using IPAQ).
Neuromuscular measures	HA: Step test; Timed up & go (TUG); Grip strength (dynamometry)	HA: chair rise; walking speed; leg rise; Grip test CAPI: self-reported steadiness	HA: TUG; GaitRite mat; Grip strength (dynamometry); heel bone ultrasound test CAPI: self-reported steadiness
Sensory measures	HA: lens photography; retinal imaging; intra-ocular pressure; auto refraction	CAPI: self-reported vision and hearing	HA: retinal imaging CAPI: self-reported vision and hearing
Cognitive system	HA: mini-mental state examination (MMSE); colour trials; Animal fluency; Montreal cognitive assessment	CAPI: individual questions	HA: MMSE
Exogenous variables: Medication; falls & fractures; Fear of falling; Steadiness; Pain; ADL	CAPI	CAPI: but does not include fear of falling	CAPI

Key: (Source: ELSA data; TILDA data; NICOLA data proposal); CAPI is Computer Assisted Personal Interview; SCQ is Self-Completed Questionnaire; HA is Health Assessment.

Appendices IX

Ethics approval confirmation RG3 form (Ulster University)

UNIVERSITY OF ULSTER **RESEARCH GOVERNANCE**

RG3 Filter Committee Report Form

Project Title	The association between physical activity and balance in an older free living adult population (50 years of older)
Chief Investigator	Professor Suzanne McDonough
Filter Committee	Institute of Nursing and Health Research

This form should be completed by Filter Committees for all research project applications in categories A to D (*for categories A, B, and D the University's own application form – RG1a and RG1b – will have been submitted; for category C, the national, or ORECNI, application form will have been submitted).

Where substantial changes are required the Filter Committee should return an application to the Chief Investigator for clarification/amendment; the Filter Committee can reject an application if it is thought to be unethical, inappropriate, incomplete or not valid/viable.

Only when satisfied that its requirements have been met in full and any amendments are complete, the Filter Committee should make one of the following recommendations:

The research proposal is complete, of an appropriate standard and is in

- category A and the study may proceed* [x]
- category B and the study must be submitted to the University's Research Ethics Committee** Please indicate briefly the reason(s) for this categorisation [ ]
- category C and the study must be submitted to ORECNI along with the necessary supporting materials from the Research Governance Section*** [ ]
- category D and the study must be submitted to the University's Research Ethics Committee** [ ]

Signed: *George Kernohan* Date: 11 January 2017
Chairperson of Filter Committee

***The application form and this assessment should now be returned to the Chief Investigator. The Filter Committee should retain a copy of the complete set of forms.**

**** The application form and this assessment should now be returned to the Chief Investigator so that he/she can submit the application to the UUREC via the Research Governance section. The Filter Committee should retain a copy of the complete set of forms for their own records.**

***** The application form and this assessment should now be returned to the Chief Investigator so that he/she can prepare for application to a NRES/ORECNI committee. The Filter Committee should retain a copy of the complete set of forms for their own records.**

<u>For all categories, details of the application and review outcome should be minuted using the agreed format and forwarded to the Research Governance section</u>

Version 3 (10/07) 1

Appendices IX (continued)

Ethics approval confirmation RG3 form (Ulster University)

Please complete the following

The application should be accompanied by an appropriate and favourable Peer Review Report Form (if not, the Filter Committee should be prepared to address this as part of its review). Please comment on the peer review (include whether or not there is evidence that the comments of the peer reviewers have been addressed).

Peer review is complete and there are no outstanding issues of serious ethical concern.

Please provide an assessment of all component parts of the application, including questionnaires, interview schedules or outline areas for group discussion/unstructured interviews.

This is a low risk study to examine physical activity and balance in older people using secondary analysis of data from large longitudinal ageing studies in Republic of Ireland, England and from N. Ireland.

Please comment on the consent form and information sheet, in particular the level of language and accessibility.

No issues in this area.

Please comment on the qualifications of the Chief and other investigators.

Well qualified investigators

Please comment on the risks present in conducting the study and whether or not they have been addressed.

This is a low risk study. Good practice in research will be observed in the management of any risks associated with data management.

Please indicate whether or not the ethical issues have been identified and addressed.

Yes, ethical issues are addressed. The benefit in new knowledge clearly outweighs any risks associated with the study.

Please comment on whether or not the subjects are appropriate to the study and the inclusion/exclusion criteria have been identified and listed

Yes, the data samples are clear.

Version 3 (10/07) 2

Appendix X
Development of the predictive model of PA and balance controlling for exogenous variables using TILDA

Mplus syntax was found for confirmatory factor analysis with continuous factor indicators in the mplus manual (Muthén & Muthén, 1998-2017). A confirmatory factor analysis was run for a one factor model for balance using the data from wave 1 in mplus (version 7.4). The fit statistics outlined in table 1 indicated a good fit.

Table 1: Fit statistics for balance at wave 1:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
Balance wave 1	183750.70	183884.61	150.55	8	0.00	0.05	(0.04, 0.05)	0.97	0.94	0.03

The process was then repeated process for balance wave 2 data as per table 2 below, and the fit statistics showed a poor fit. Upon examination of the modification indices it was found that values for vision and hearing were high (476.413; 476.410).

Table 2: Fit statistics for balance at wave 2:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
w2	187353.66	187477.55	684.72	9	0.00	0.10	(0.09, 0.12)	0.88	0.81	0.05

Then, using the same process a 2-factor model for balance at wave 1 and 2 simultaneously without any restrictions (configural invariance) was run. Fit statistics (table 3) were examined to assess the fit of the proposed model. The fit statistics, as expected, due to the discovery of the high value for vision and hearing seen in wave 2, indicated that the model was not a good fit. The model modification indices were then examined, and it was found that the two residuals were particularly correlated with high modification index scores: hearing at wave 1 (HEAR02w1) with vision at wave 1 (VIS01W1) (271.732) and hearing at wave 2 (HEAR02W2) with vision at wave 2 (VIS01W2) (240.418), indicating a potentially additional source of variance within the model.

Appendix X (continued)

Development of the predictive model of PA and balance controlling for exogenous variables using TILDA

Table 3: Model 1: Confirmatory factor model for balance at wave 1 and wave 2:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
Model 1	339954.70	340255.95	827.489	47	0.0000	0.05	(0.04, 0.05)	0.88	0.83	0.05

Thus, a two-factor model for balance at wave 1 and wave 2 with the 2 correlated residuals (HEAR02w1 with VIS01W1 and HEAR02W2 with VIS01W2) was then re-run in mplus. The parameter estimates were allowed to be freely estimated. The fit statistics outlined in table 4 were again analysed to assess the fit of the model. The fit indices indicated that the model fit had improved, and upon examining the model modification indices, it was found that the modification index scores of vision at wave 2 (VIS01W2) WITH hearing at wave 1 (HEAR02W1) (42.536) and hearing at wave 2 (HEAR02W2) WITH vision at wave 1 (VIS01W1) (58.894) were high.

Table 4: Model 1: Two factor model for balance at wave 1 and wave 2 with 2 correlated residuals:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
Correlated residuals (x2)	339178.75	339494.00	375.48	45	0.0000	0.03	(0.03, 0.03)	0.95	0.93	0.04

Thus, as both residuals are self-reported measures of sensory system health, it was decided to correlate them and re-run the model. The fit indices outlined in the above table indicate that correlating the four residuals (HEAR02W1 WITH VIS01W1; HEAR02W2 WITH VIS01W2; HEAR02W2 WITH VIS01W1; and VIS01W2 WITH HEAR02W1) further improved the model fit. However, the modification statistics indicate that steadiness and balance have a correlated relationship

Appendix X (continued)

Development of the predictive model of PA and balance controlling for exogenous variables using TILDA

where the "ON/BY Statements" show a modification index of 22.471 between steadiness in wave 1 and balance in wave 2: STEAD1_S ON BALANCE2 /BALANCE2 BY STEAD1_S.

Table 5: Model 1: Two factor model for balance at wave 1 and wave 2 with 4 correlated residuals:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
correlated residuals (x4)	338974.21	339303.48	253.90	43	0.0000	0.03	(0.02, 0.03)	0.97	0.95	0.03

All the self-reported measures in the model: self-reported vision, self-reported hearing, and self-reported steadiness may potentially be causing a 'method effect' where self-reported measures may be biased due to recall issues, or over or under estimation of effect. In addition, the self-reported measure of steadiness is comprised of three questions that were summed for this model, and which haven't been validated. Thus, there was a decision to be made on how to deal with these issues moving forward. There is an argument to not allow for these variances in the model as we need to model other databases and therefore we should avoid over complicating the model by for example, adding additional parameters. As a result, it was decided to not include an additional correlation of 'method effect' in the model at this stage. The next step was to look at whether the observed variables behaved the same across wave 1 and wave 2 using measurement invariance, which involved restricting the factor loadings of each observed variable within the model. The fit statistics suggested that the model was a good fit as outlined in the table below.

Appendix X (continued)
Development of the predictive model of PA and balance controlling for exogenous variables using TILDA

Table 6: Model 2: Two factor model for balance at wave 1 and 2 analysed using metric invariance:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
Metric invariance	339283.97	339571.20	342.29	49	0.0000	0.03	(0.02, 0.03)	0.96	0.94	0.05

Further analysis of the "variances/residual variance" modification indices showed high values for the unobserved measure of balance at wave 1 (Balance1) and 2 (Balance2) with values of 81.410 and 81.414 respectively due to the restriction imposed on the factor (balance) for the purposes of identification. The fit indices suggest the model was working well and thus, the model was further investigated using scalar invariance analysis, where both the loadings and intercepts are constrained to be equal. Scalar invariance analysis explores whether the observed variables are performing in the same way at wave 1 and 2. The fit statistics found that the model fit had declined and on exploration of the modification indices (means/intercepts/threshold) it was found that balance at wave 1 and 2 were high (Balance1 was 87.421, and Balance2 was 87.418). In addition, the modification indices showed high results for Grip test (e.g. 332.416), MMSE (109.305) and Hearing (62.514). Thus, restrictions on the intercepts for the observed variables of Grip test, MMSE and Hearing were removed for wave 1 and wave 2 and because the factor of the means had been introduced it was possible to remove the restriction on balance for the purposes of the identification and instead use TUG as the means of model identification. TUG was selected because it had previously performed well across both wave 1 and wave 2. Thus, the model was re-run and the following fit statistics suggest that the model now provides a good explanation of the data:

Appendix X (continued)
Development of the predictive model of PA and balance controlling for exogenous variables using TILDA

Table 7: Two factor model for balance at wave 1 and 2 analysed using scalar invariance:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
Model 3: Scalar invariance	339117.30	339397.53	313.07	50	0.0000	0.03	(0.02, 0.03)	0.96	0.95	0.04

Having now corrected for measurement error within the model, physical activity was introduced. Firstly, a factor loading of 0.76 was imposed on physical activity as this was the reliability value highlighted by Craig et al., 2003. The model was re-run and the following represent the fit statistics:

Table 8: Two factor model for balance at wave 1 and 2 including physical activity:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
Model 4: balance and physical activity	414149.79	4144493.07	593.99	70	0.0000	0.03	(0.03, 0.03)	0.94	0.92	0.04

The fit statistics showed that the inclusion of PA within the model resulted in a poorer fit. Further investigation of the model results showed that the same measures are being used at both waves 1 and 2; the ‘standardised model’ results showed: TUG (0.71) and Steadiness (0.70) are highly related to balance but that vision (0.30) and hearing (0.25) are not; that the rank of ordering has been maintained between balance at wave 2 and balance at wave 1 (0.99); that although the relationship between balance at wave 2 and PA at wave 1 is not significant (0.02) it is in the expected direction; that PA at wave 2 is affected by PA at wave 1 (0.44); that PA at wave 2 is affected by balance at wave 1 (-0.17); that balance at wave 1 affects PA at wave 2 (-0.18); that balance and PA at wave 1 are highly correlated (-0.36). Exploration of the standardised intercepts (STD standardization) showed that whilst MMSE and vision were different, this was a marginal difference, but that grip at wave 1 and 2 were not equal (25.96 and 29.75 respectively). Additionally, the model modifications indices showed that whilst grip is a component of balance there is a lot of residual variance in grip, and also that PA is highly related to grip (where on statements show that PA on grip at wave 1 is 194.58 and wave 2 is 147.36). As a result, grip was correlated with PA and the model was rerun showing the following fit statistics:

Table 9: Two factor model for balance and physical activity including a correlation of grip:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
Model 4: balance and physical activity	413812.58	414176.88	401.03	67	0.0000	0.03	(0.02, 0.03)	0.96	0.95	0.03

The model now appears to adequately describe the data. The next step was to incorporate the exogenous variables into the model. The exogenous variables described in the table were prepared as described using SPSS. Then in mplus, the internal consistency of the multiple item measures was assessed using a 3-factor model. A factor model was run for pain, sleep and alcohol as these covariates had multiple items. The results showed that the correlation between the two items in pain was not good, and upon investigation it was found that one of the items related more to medication for pain rather than pain itself and so it was decided to include only one item for pain (pain1w1). The correlation on the three items for

Appendix X (continued)
Development of the predictive model of PA and balance controlling for exogenous variables using TILDA

sleep was not good, so it was decided that all three items would remain separate. The correlation was good for alcohol with a factor loading of 0.929 indicating a strong internal consistency for alcohol. The exogenous variables were then introduced into the model one by one (re_sex; re_meds; re_falls; Edu_prime; Edu_second (Edu_third was excluded so that it could be the comparison group); re_age; pain1w1; alch1_sum; sleep2w1; sleep3w1; re_fefall; re_adlw1. The fit statistics were checked at each step and where necessary the modification indices were checked, and correlations introduced to improve model fit. The following correlations were introduced: sex on grip at wave 1 and 2 (research suggests that men have a stronger grip than women); age on balance at wave 1 (age is a key risk factor for balance); age on grip at wave 1 and 2 (age affects grip strength); age on steadiness (age affects balance); fear of falling on balance at wave 1 (fear of falling is a risk factor for balance); medication on balance at wave 1 (medication is a risk factor for balance); education on balance at wave one (low SES is a risk factor for balance); MMSE on education (cognitive ability affects SES); pain on balance at wave 1 (pain is a risk factor for balance); alcohol on balance at wave 1 (alcohol is a risk factor for balance); ADL on balance at wave 1 (disability in any ADL is a risk factor for balance). The model was re-run and the fit statistics indicate that the model reflects the data well:

Table 10: Two factor model for balance and physical activity including exogenous variables:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	P	Estimate	90 % C.I.	CFI	TLI	Value
Model 5: Final model	90876.18	91322.57	503.74	205	0.0000	0.03	(0.03, 0.03)	0.95	0.94	0.04

Appendix XI
M-plus input file for final model of physical activity and balance using data from the TILDA study

DATA: FILE IS TILDA 1 wave 1 & 2 MERGED variables_covariates_3.dat;
define: IPA_1 = IPAQmmw1/1000;
define: IPA_2 = IPAQmmw2/1000;
VARIABLE: NAMES ARE
id houshold cluster hcweight stratum capiweig in_scq scqweigh
age sex Educ_lev ill31w1 ill41w1 ill31w2 ill41w2 Vis01w1 Vis02w1 Vis03w1
Hear01w1 Hear02w1 DISvis1 DIShear1Vis01w2 Hear02w2 Hear03w2
Hear04w2 Hear05w2Fall01w1 Fall02w1 Fall03w1Fall04w1 Fall05w1Fall06w1
Fall07w1 Fall08w1 Fall09w1 Fall01w2 Fall02w2 Fall03w2 Fall04w2 Fall05w2
Fall09w2 Stead1w1 Stead2w1 Stead3w1 Stead1w2 Stead2w2 Stead3w2
Pain1w1 Pain2w1 Pain1w2 Pain2w2 ADL1w1 ADL2w1 ADL3w1 ADL4w1
ADL5w1 ADL6w1 ADL7w1 ADL8w1 ADL9w1 ADL10w1 ADL11w1 ADL12w1
ADL13w1 ADL14w1 ADL15w1 ADL16w1 ADL17w1 fl025_1 fl025_2 fl025_3
fl025_4 fl025_5 fl025_6 fl025_7 fl025_8ADL1w2 ADL2w2 ADL3w2 ADL4w2
ADL5w2 ADL6w2 ADL7w2 ADL8w2 ADL9w2 ADL10w2 ADL11w2 IPAQv1w1
IPAQv2w1 IPAQv3w1 IPAQm1w1 IPAQm2w1 IPAQm3w1 IPAQw1w1
IPAQw2w1 IPAQw3w1 IPAQs1w1 IPAQs2w1 IPAQmmw1 IPAQPAw1
IPAQv1w2 IPAQm1w2 IPAQw3w2 IPAQm3w2 IPAQv3w2 walkingmet2
moderatemet2 vigorousmet2 IPAQmmw2 IPAQPAw2 Sleep1w1 Sleep2w1
Sleep3w1 Sleep1w2 Sleep2w2 Sleep3w3 ha_weight in_ha
R_Height_Centimetres MMSEw1 MMSEw2 FRbmi GripDw1 GripNDw1
GripDw2 GripNDw2 TUGsw1 TUGs1w1 TUGsw2 cage1w1 cage2w1 cage3w1
cage1w2 cage2w2 cage3w3 NADLw1 NIADLw1 NADLw2 NIADLw2 PolyMDw1
PolyMDw2 INCASSas SES age2 gd002 edu_level Stead1_sum Stead2_sum
sleep_sum pain_sum alch1_sum Edu_prime Edu_second Edu_third re_sex
re_fefall re_falls MDmeds4 re_meds re_ses re_age re_METmins re_ADLw1;
missing are GripDw1 (-999, 98, 99, -1) TUGsw1 (-999, 98, 99, -1)
Stead1_sum (-999, 98, 99, -1) Vis01w1 (-999, 98, 99, -1) Hear02w1 (-999, 98,
99, -1) MMSEw1(-999, 98, 99, -1) GripDw2 (-999, 98, 99, -1) TUGsw2 (-999, 98,
99, -1) Stead2_sum (-999, 98, 99, -1) Vis01w2 (-999, 98, 99, -1) Hear02w2 (-
999, 98, 99, -1) MMSEw2(-999, 98, 99, -1) stratum (-999, 98, 99, -1) capiweig (-
999, 98, 99, -1) cluster (-999, 98, 99, -1) re_sex (-999, 98, 99, -1) re_age (-999,
98, 99) re_meds (-999, 98, 99, -1) re_falls (-999, 98, 99, -1) Edu_prime (-999,
98, 99, -1) Edu_second (-999, 98, 99, -1) Pain1w1 (-999, 98, 99, -1) alch1_sum
(-999, 98, 99, -1) Sleep2w1 (-999, 98, 99, -1) Sleep3w1(-999, 98, 99, -1)
re_fefall (-999, 98, 99, -1) re_METmins (-999, 98, 99, -1) re_ADLw1 (-
999,98,99, -1);
subpopulation is (age ge 50);
USEVAR are GripDw1 TUGsw1 Stead1_sum Vis01w1 Hear02w1 MMSEw1
GripDw2 TUGsw2 Stead2_sum Vis01w2 Hear02w2 MMSEw2 re_sex re_meds
re_falls Edu_prime Edu_second re_age Pain1w1 alch1_sum Sleep2w1

Appendix XI (continued)
M-plus input file for final model of physical activity and balance using data from the TILDA study

```
Sleep3w1  re_fefall re_ADLw1 stratum capiweig cluster IPA_1 IPA_2  ;
cluster = cluster; stratification = stratum; weight = capiweig;
ANALYSIS: TYPE = COMPLEX;
ANALYSIS: estimator = mlr;
iterations = 5000;
MODEL: Balance1 BY GripDw1* (1);
Balance1 BY TUGsw1@1; !(2);
Balance1 BY Stead1_sum (3) ;
Balance1 BY Vis01w1 (4);
Balance1 BY Hear02w1 (5);
Balance1 BY MMSEw1 (6);
IPA1 by IPA_1@0.872; !rel = 0.76
IPA_1@2.863; ! residual variance = (1 - reliability)*sample variance[11.927]
[GripDw1];
[TUGsw1];
[Stead1_sum];
[Vis01w1];
[Hear02w1];
[MMSEw1] (12);
MODEL: Balance2 BY GripDw2* (1);
Balance2 BY TUGsw2@1;! (2);
Balance2 BY Stead2_sum (3);
Balance2 BY Vis01w2 (4);
Balance2 BY Hear02w2 (5);
Balance2 BY MMSEw2 (6);
[GripDw2];
[TUGsw2];
[Stead2_sum];
[Vis01w2];
[Hear02w2];
[MMSEw2] (12);
GripDw1 with GripDw2;
TUGsw1 with TUGsw2;
Stead1_sum with Stead2_sum;
Vis01w1 with Vis01w2;
Hear02w1 with Hear02w2;
MMSEw1 with MMSEw2;
HEAR02W1 WITH VIS01W1;
HEAR02W2 WITH VIS01W2;
VIS01W2 WITH HEAR02W1;
HEAR02W2 WITH VIS01W1;
Balance2 on Balance1;
```

Appendix XI (continued)
M-plus input file for final model of physical activity and balance using data from the TILDA study

```
IPA_2 on IPA1;
Balance1 on IPA1;
Balance2 on IPA1;
IPA_2 on Balance1;
Balance2 on IPA_2;
GRIPDW1 ON IPA1;
GRIPDW2 ON IPA1;
IPA1 on re_sex re_age
re_meds
re_falls
Edu_prime Edu_second
Pain1w1
alch1_sum
Sleep2w1 Sleep3w1
re_fefall
re_ADLw1;
GRIPDW1 ON RE_SEX;
GRIPDW2 ON re_sex;
BALANCE1 ON RE_AGE;
GRIPDW2 ON RE_AGE;
GRIPDW1 ON RE_AGE;
STEAD1_SUM ON RE_AGE;
STEAD2_SUM ON RE_AGE;
BALANCE1 ON RE_FEFALL;
BALANCE1 ON RE_MEDS;
BALANCE1 ON EDU_PRIME;
BALANCE1 ON EDU_second;
MMSEW1   ON EDU_PRIME;
MMSEW1 ON EDU_SECOND;
MMSEW2   ON EDU_PRIME;
MMSEW2 ON EDU_SECOND;
BALANCE1 ON PAIN1W1;
BALANCE1 ON ALCH1_SUM;
BALANCE1 ON RE_ADLW1;
Model indirect:
Balance2 ind Balance1 IPA1;
Balance2 ind IPA1;
Balance1 ind re_sex;
OUTPUT: sampstat STANDARDIZED MODINDICES (all);
output: sampstat;
```

Appendix XII
M-plus input file for development of Latent class analysis of physical activity measure from the ELSA study

```
TITLE: Latent class analysis using the physical activity measure from the ELSA
study
Data: file is elsa lcg (2).dat;
define: cff_c1 = cffc1/5;
define: cff_c2 = cffc2/5;
define: cff_c4 = cffc4/5;
define: cff_c5 = cffc5/5;
define: GRIPa_2 = GRIPa2/20;
define: GRIPb_2 = GRIPb2/20;
define: GRIPc_2 = GRIPc2/20;
define: GRIPa_4 = GRIPa4/20;
define: GRIPb_4 = GRIPb4/20;
define: GRIPc_4 = GRIPc4/20;
define: GRIPa_6 = GRIPa6/20;
define: GRIPb_6 = GRIPb6/20;
define: GRIPc_6 = GRIPc6/20;
VARIABLE: NAMES ARE
idauniq Eyea1 Eyea2 Eyea3 Eyea4 Eyea5 Eyea6 Eyeb1 Eyeb2 Eyeb3 Eyeb4
Eyeb5 Eyeb6 Eyec1 Eyec2 Eyec3 Eyec4 Eyec5 Eyec6 Eara1 Eara2 Eara3
Eara4 Eara5 Eara6 Earb1 Earb2 Earb3 Earb4 Earb5 Earb6 Steada1 Steada2
Steada4 Steadb1 Steadb2 Steadb4 Gaita1 Gaita2 Gaita3 Gaita4 Gaita5 Gaita6
Gaitb1 Gaitb2 Gaitb3 Gaitb4 Gaitb5 Gaitb6 cfra1 cfra2 cfra4 cfra5 cfra6 cfrb1
cfrb2 cfrb4 cfrb5 cfrb6 cffc1 cffc2 cffc4 cffc5 cfmd1 cfmd2 cfmd4 cfmd5 cfre1
cfre2 cfre4 cfre5 cfre6 GRIPa2 Gripa4 Gripa6 GRIPb2 Gripb4 Gripb6 GRIPc2
Gripc4 Gripc6 Stat2 stat4 stat6 Stbt2 stbt4 stbt6 Stct2 stct4 stct6 Legt2 legt4
legt6 Legst2 legst4 legst6 Mcht2 mcht4 mcht6 w4nurwt w3lwgt dhager3 w5lwgt
w5xwgt w5scwt w6lwgt w6xwgt w6scwt Earb_1 Earb_2 Earb_3 Earb_4 Earb_5
Earb_6 idahhw1 w1wgt dhager1 ahsecls21 astratif1 idahhw2 sampsta2 dhager2
hseclst2 astratif2 w2wgt scw2wgt idaindw2 w2wtnur idahhw3 lwgt3 idahhw4
w4xwgt w4lwgt w4scwt idahhw5 indager5 idahhw6 indager6 w6nurwt Gait_a1
Gait_a2 Gait_a3 Gait_a4 Gait_a5 Gait_a6 Gait_b1 Gait_b2 Gait_b3 Gait_b4
Gait_b5 Gait_b6 Gait_a2c Gait_a3c Gait_a4c Gait_a5c Gait_a6c Gait_b2c
Gait_b3c Gait_b4c Gait_b5c Gait_b6c mcht2c mcht4c mcht6c m2 m4 m6 m2c
m4c m6c aGait1 aGait2 aGait3 aGait4 aGait5 aGait6 bGait1 bGait2 bGait3
bGait4 bGait5 bGait6 disex dimar1 ADLa1 ADLa2 ADLa3 ADLa4 ADLa5 ADLa6
Falla1 Falla2 Falla3 Falla4 Falla5 Falla6 Paina1 Paina2
 Paina3 Paina4 Paina5 Paina6 EDU1 EDU2 EDU3 EDU4 EDU5
 Alca2 Alca3 Alca5 Alca4 Alca6 PAa1 PAb1 PAc1
 PAa2 PAb2 PAc2 PAa3 PAb3 PAc3 PAa4 PAb4 PAc4 PAa5
 PAb5 PAc5 PAa6 PAb6 PAc6 PAsum2 PAsum3 PAsum5
 ADLi3 sleepa3 Sleepa4 Sleepe4 sleepa6 sleepe6 worka1
 workb1 workc1 workd1 worke1 workf1 worka2
```

Appendix XII (continued)
M-plus input file for development of Latent class analysis of physical activity measure from the ELSA study

```
workb2 workc2 workd2 worke2 workf2 worka4 workb4 workc4 workd4 worke4
workf4 workg4 worka5 workb5 workc5 workd5 worke5 workf5 workg5 worka6
workb6 workc6 workd6 worke6 workf6 workg6 indager4 age_cat Agec_1
Agec_2 Agec_3 Agec_4 Agec_5 Agec_6 VPA1 MPA1 LPA1 PAS1 VPA_1
MPA_1 LPA_1 Dem Hamer PP14 PP24 PP34 PP44 PALC4 Class4 PP13 PP23
PP33 Class3 SEX cff_c1 cff_c2 cff_c4 cff_c5 GRIPa_2 GRIPb_2 GRIPc_2
GRIPa_4 GRIPb_4 GRIPc_4 GRIPa_6 GRIPb_6 GRIPc_6;
USEVARIABLES ARE PAa1 PAb1 PAc1 Gait_a2;
missing are all (-999, -9, -8, -1);
classes = c (3);
categorical = PAa1 PAb1 PAc1;
auxiliary = Gait_a2 (BCH);! (ANOVA structure -taking into account measurement
error by a weighting strategy Bakk & Vermunt 2014)
idvariable = idauniq;
analysis: type = mixture;
savedata: file is LCA3_BCH.sav;
save is cprob;
plot: Type is plot1 plot2 plot3;
Series is PAa1(1) PAb1(2) PAc1(3);
output: tech1 tech8 tech10 tech11 tech14;
```

Appendix XIII
Development of the model of balance (ELSA)

Variables across the six waves of ELSA were firstly collated into a single SPSS file from the multiple files provided by the ELSA project team and renamed. The variables were then recoded, and scores reversed so that they were all in the same direction to enable an easier interpretation of the results.

Firstly, variables relating to balance were included in the model on a wave by wave basis, and then based on fit statistics and modification indices, the appropriate adjustments were made. When changes improved the model fit indices subsequent waves were added and the same process of iteration was followed. The following outlines the modification indices and the modification made at each wave for balance.

Step one. Configural analysis

Mplus syntax for confirmatory factor analysis with continuous factor indicators was used in mplus (Muthén & Muthén, 1998-2017) and fit statistics indicated that one residual correlation was needed: CFRE1 with CFRB1 (Cognitive measure). This was a reasonable adjustment as both questions relate to learning and recall where the same ten words were immediately recalled and then recalled after a period of time during the interview.

Table 1: Fit statistics for the latent construct of balance at wave 1

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	p value	Estimate	(90 % C.I.)	CFI	TLI	Value
Post-mod	224924.36	225207.02	236.17	49	0.0000	0.99	(0.02, 0.03)	0.99	0.99	0.02

Wave two data was then added to the model, but fit statistics had declined and so the following changes were made based on the modification indices and the two factor model for balance wave one and two is shown below where 12 additional residual correlations were introduced (Eye2 with Eye; Ear2 with Ear; Steady2 with Steady; Cog 2 with Cog; CFF_C2; Eyeb2 with Eyeb1; Wara2 with Eara1; Earb2 with Earb1; Steadb2 with Steadb1; Cfra2 with Cfra1; Cfre2 with Cfre1; Cfre2 with Cfrb2). Line 1 shows fit statistics pre- modification and line 2 shows fit statistics post-modifications.

Appendix XIII (continued)
Development of the model of balance (ELSA)

Table 2: Fit statistics for balance at wave 1 and 2:

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	p value	Estimate	(90 % C.I.)	CFI	TLI	Value
Pre-mod	434298.64	435002.55	9047.77	393	0.0000	0.06	(0.05, 0.06)	0.87	0.85	0.09
Post-mod	425501.82	426288.55	1691.32	381	0.0000	0.02	(0.02, 0.02)	0.98	0.99	0.04

Wave three data for balance was then added to the model, but fit statistics had declined and so based on modification indices the following eleven residual correlations were added to the model (Eye3 with Eye; Eye3 with Eye2; Ear3 with Ear1; Gait3 with Gait2; Eyeb3 with Eyeb2; Eyec3 with Eyec2; Eara3 with Eara2; Eara3 with Eara1; Earb3 with Earb1; Earb3 with Earb2). Fit statics for a three-factor model of balance is shown in Table 3 where line 1 shows fit statistics pre- modification and line 2 shows fit statistics post-modifications.

Table 3: Fit statics for balance at wave 1, 2 and 3

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	p value	Estimate	(90 % C.I.)	CFI	TLI	Value
Pre-mod	505338.18	505320.03	6957.02	600	0.0000	0.04	(0.04, 0.04)	0.92	0.91	0.09
Post-mod	498647.47	499699.45	2267.98	590	0.0000	0.02	(0.02, 0.02)	0.98	0.98	0.05

Wave four data was then added to the model and the same process as outlined above was followed. As a result, the following 25 residual correlations were added (Eye 4 with Eye3; Ear4 with Ear3; Gait4 with Gait3; Chair4 with Chair 2; Grip4 with Grip2; Cog4 with Cog1; Cog4 with Cog2; Stady4 with Stady2; Eyec4 with Eyec3; Eara4 with Eara3; Earb4 with Earb3; Steadb4 with Steadb2; Cfre4 with Cfrb4; Cff-c4 with Cff-c1; Cff-c4 with Cff-c2; Eye4 with Eye; Eye4 with Eye2; Ear4 with Ear; Ear4 with Ear2; Steady4 with Steady1; Eara4 with Eara1; Eara4 with Eara2; Earb4 with Earb2; Steada4 with Steada1; Steadb4 with Steadb1). Table 4 shows the fit statistics where line 1 shows fit statistics pre- modification and line 2 shows fit statistics post-modifications.

Appendix XIII (continued)
Development of the model of balance (ELSA)

Table 4: Fit statistics for balance at wave 1, 2, 3, 4

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	p value	Estimate	(90 % C.I.)	CFI	TLI	Value
Pre-mod	723558.70	725084.92	15044.81	1382	0.0000	0.03	(0.03, 0.03)	0.89	0.89	0.09
Post-mod	710796.57	712501.92	24636.00	1357	0.0000	0.02	(0.02, 0.02)	0.98	0.97	0.04

Then, using the same process wave five data for balance was added to the model and resulted in 24 additional Eye5 with Eye4; Ear5 with Ear4; Gait5 with Gait4; Cog5 with Cog4; Eyea5 with Eyea4; Eyeb5 with Eyeb4; Eara5 with Eara4; Earb5 with Earb4; Cfra5 with Cfra4; Xfmd5 with Cfmd4; Cfre5 with Cfre4; Cfre5 with Cfrb5; Cff-c5 with Cff-c4; Eye5 with Eye3; Ear5 with Ear3; Gait5 with Gait3; Cog5 with Cog; Cog5 with Cog2; Earb4 with Earb1; Eyeb5 with Eyeb4; Eara5 with Eara3; Earb5 with Earb3; Cff-c5 with Cff-c1; Cff-c5 with Cff-c2). Table 5 shows the fit statistics where line 1 shows fit statistics pre-modification and line 2 shows fit statistics post-modifications.

Table 5: Fit statistics for balance waves 1, 2, 3, 4, and 5

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	p value	Estimate	(90 % C.I.)	CFI	TLI	Value
Pre-mod	877063.41	879085.92	15754.79	2063	0.0000	0.03	(0.03, 0.03)	0.92	0.91	0.08
Post-mod	868345.70	870525.99	8665.77	2041	0.0000	0.02	(0.02, 0.02)	0.96	0.96	0.06

Wave six data was then added to the model and again where appropriate residual correlations were added to the model. Table 6 shows the fit statistics for a six-factor model of balance pre-modifications (line 1). Line 2 shows fit statistics post an additional 35 residual correlations (Eye6 with Eye5; Ear6 with Ear2; Ear6 with Ear4; Ear6 with Ear5; Gait6 with Gait5; Chair6 with Chair4; Grip6 with Grip2; Grip6 with Grip4; Cog6 with Cog5; Eyea6 with Eyea5; Eyeb6 with Eyeb5; Eyec6 with Eyec5; Eara6 with Eara5; Earb6 with Earb4; Earb6 with Earb5; Cfra6 with Cfra5; Cfre6 with Cfre5; Eye5 with Eye2; Ear5 with Ear2; Eye6 with Eye4; Ear6 with Ear; Chair6 with Chair2; Cog6 with Cog2; Ear5 with Ear; Eye6 with Eye3; Cog6 with Cog; Ear with Eye; Ear2 with Eye2; Ear3 with Eye3; Ear4 with Eye4; Ear5 with Eye5; Ear6 with Eye6). A decision was made not to correlate M6c with M4c as this measure has only one item; and a correlation was introduced between Ear and Eye as both measures are measures of the sensory system and may be biased due to a method effect as both are also self-reported measures.

Appendix XIII (continued)
Development of the model of balance (ELSA)

Table 6: Fit statistics for balance waves 1, 2, 3, 4, 5, 6

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
	Akaike (AIC)	Bayesian (BIC)	value	df	p value	Estimate	(90 % C.I.)	CFI	TLI	Value
Pre-mod	1014547.15	1017122.40	22704.70	3046	0.0000	0.03	(0.03, 0.03)	0.91	0.91	0.08
Post-mod	998102.48	1000930.92	8756.06	3011	0.0000	0.01	(0.01, 0.01)	0.98	0.97	0.05

Step 2: Metric invariance

A series of successive restrictions on the factor loadings on each measure of balance across all six waves was imposed and the fit statistics met the criteria.

Table 7: Fit statistics for balance waves 1, 2, 3, 4, 5, 6 with restrictions on factor loadings (metric invariance)

Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
Akaike (AIC)	Bayesian (BIC)	value	df	p value	Estimate	(90 % C.I.)	CFI	TLI	Value
998232.46	1000778.78	8787.34	3050	0.0000	0.01	(0.01, 0.01)	0.98	0.97	0.05

Step 3: Scalar invariance

A series of successive restrictions were then placed on the intercepts on each measure of balance across all six waves and Table 7, line 1 shows that the fit statistics had deteriorated significantly, and so restrictions were removed using an iterative process (where line 2 shows the removal of restrictions on Chair rise test (M2C and M6C) as well as Hand grip test (Gripa2, Gripb2; Gripc2; Gripa6, Gripb6, Gripc6) at waves two and six; Line 3 shows the additional removal of restriction on Ear at wave one and two (Eara1, Earb1, Eara2, Earb2); Line 4 shows the removal of restrictions on Steadiness at time 4 (Steada4, Steadb4); Line 5 shows removal of restrictions on Gait test at waves two, three, five and six (Gait-a2c; Gait-b2c; Gait-a6c; Gaitb6c; Gait-a3c; Gait-b3c; Gait-a5c, Gait-b5c); and line 6 shows restrictions on the second order construct of balance (Balance, Balance2,Balance3, Balance4, Balance5, Balance6).

Appendix XIII (continued)
Development of the model of balance (ELSA)

Table 7: Scalar invariance waves 1, 2, 3, 4, 5, 6

	Information Criteria		Chi squared			RMSEA		CFI/TLI		SRMR
Line	Akaike (AIC)	Bayesian (BIC)	value	df	p value	Estimate	(90 % C.I.)	CFI	TLI	Value
1	1007736.07	1009837.89	16911.03	3112	0.0000	0.02	(0.02, 0.02)	0.94	0.94	0.02
2	999543.83	1001699.53	9998.40	3104	0.0000	0.02	(0.01, 0.02)	0.97	0.97	0.05
3	999241.29	1001425.91	9731.51	3100	0.0000	0.01	(0.01, 0.02)	0.97	0.97	0.05
4	999107.31	1001306.41	9611.95	3098	0.0000	0.01	(0.01, 0.02)	0.97	0.97	0.05
5	998664.83	1000914.55	9245.05	3091	0.0000	0.01	(0.01, 0.01)	0.97	0.97	0.05
6	998383.88	1000677.02	8986.99	3085	0.0000	0.01	(0.01, 0.01)	0.97	0.97	0.05

Appendix XIV
M-plus input file for model of balance using measures from ELSA

```
TITLE: ELSA model of balance across wave 1,2,3,4,5 & 6
Data: file is elsa w1 and 2 3 4 5 6 merged_6.dat;
define: cff_c1 = cffc1/5;
define: cff_c2 = cffc2/5;
define: cff_c4 = cffc4/5;
define: cff_c5 = cffc5/5;
define: GRIPa_2 = GRIPa2/20;
define: GRIPb_2 = GRIPb2/20;
define: GRIPc_2 = GRIPc2/20;
define: GRIPa_4 = GRIPa4/20;
define: GRIPb_4 = GRIPb4/20;
define: GRIPc_4 = GRIPc4/20;
define: GRIPa_6 = GRIPa6/20;
define: GRIPb_6 = GRIPb6/20;
define: GRIPc_6 = GRIPc6/20;
VARIABLE: NAMES ARE Idauniq …
USEVARIABLES ARE
Eyea1 Eyeb1 Eyec1Eara1 Earb1 Steada1 Steadb1 cfra1 cfrb1 cfmd1 cfre1
!Time 2 AND 3
Eyea2 Eyeb2 Eyec2 Eara2 Earb2 Steada2 Steadb2 cfra2 cfrb2 cfmd2 cfre2
Gait_a2c Gait_b2c Gait_a3c Gait_b3c m2c Eyea3 Eyeb3 Eyec3 Eara3 Earb3
!time4
Eyea4 Eyeb4 Eyec4 Eara4 Earb4 Steada4 Steadb4 Gait_a4c Gait_b4c
cfra4 cfrb4 cfmd4 cfre4 m4c
!time 5
Eyea5 Eyeb5 Eyec5 Eara5 Earb5 cfra5 cfrb5 cfmd5 cfre5 Gait_a5c Gait_b5c
!Time 6
Eyea6 Eyeb6 Eyec6 Eara6 Earb6 cfra6 cfrb6 cfre6 Gait_a6c Gait_b6c M6c
cff_c1 cff_c2 cff_c4 cff_c5 GRIPa_2 GRIPb_2 GRIPc_2 GRIPa_4 GRIPb_4
GRIPc_4 GRIPa_6 GRIPb_6 GRIPc_6;
missing are all (-999, -9, -8, -1);
ANALYSIS: estimator = mlr;
MODEL:
!Time 1
eye by Eyea1 @1;
eye by Eyeb1 (2);
eye by Eyec1 (3);
[Eyea1] (20);
[Eyeb1] (21);
[Eyec1] (22);
ear by Eara1 @1;
ear by Earb1 (4);
```

Appendix XIV (continued)
M-plus input file for model of balance using measures from ELSA

```
[Eara1];
[Earb1];
steady by Steada1 @1;
steady by Steadb1 (5);
[Steada1] (25);
[Steadb1] (26);
cog by cfra1 @1;
cog by cfrb1 (6);
cog by cff_c1 (7);
cog by cfmd1 (8);
cog by cfre1 (9);
[cfra1] (27);
[cfrb1] (28);
[cff_c1] (29);
[cfmd1] (30);
[cfre1] (31);
!Time 2
eye2 by Eyea2 @1;
eye2 by Eyeb2 (2);
eye2 by Eyec2 (3);
[Eyea2] (20);
[Eyeb2] (21) ;
[Eyec2] (22);
ear2 by Eara2 @1;
ear2 by Earb2 (4);
[Eara2];
[Earb2] ;
Steady2 by Steada2 @1;
Steady2 by Steadb2 (5);
[Steada2] (25);
[Steadb2] (26);
cog2 by cfra2 @1;
cog2 by cfrb2 (6);
cog2 by cff_c2 (7);
cog2 by cfmd2 (8);
cog2 by cfre2 (9);
[cfra2] (27);
[cfrb2] (28);
[cff_c2] (29);
```

Appendix XIV (continued)
M-plus input file for model of balance using measures from ELSA

```
[cfmd2] (30);
[cfre2] (31);
Gait2 by Gait_a2c @1;
Gait2 by Gait_b2c (10);
[Gait_a2c];
[Gait_b2c];
GRIP2 by GRIPa_2 @1;
Grip2 by GRIPb_2 (11);
Grip2 by GRIPc_2 (12);
[GRIPa_2];
[GRIPb_2];
[GRIPc_2];
Chair2 by M2c @1;
M2c@0;
[M2C];
! Time 3
eye3 by Eyea3 @1;
eye3 by Eyeb3 (2);
eye3 by Eyec3 (3);
 [Eyea3] (20);
[Eyeb3] (21);
[Eyec3] (22);
ear3 by Eara3 @1;
ear3 by Earb3 (4);
[Eara3] (23);
[Earb3] (24);
Gait3 by Gait_a3c @1;
Gait3 by Gait_b3c (10);
[Gait_a3c];
[Gait_b3c];
!Time 4
eye4 by Eyea4 @1;
eye4 by Eyeb4 (2);
eye4 by Eyec4 (3);
[Eyea4] (20);
[Eyeb4] (21);
[Eyec4] (22);
ear4 by Eara4 @1;
ear4 by Earb4 (4);
```

Appendix XIV (continued)
M-plus input file for model of balance using measures from ELSA

```
[Eara4] (23);
[Earb4] (24);
Gait4 by Gait_a4c @1;
Gait4 by Gait_b4c (10);
[Gait_a4c] (32);
![Gait_b4c] (33);
Chair4 by M4c @1 ;
M4c@0;
[m4c] (37);
GRIP4 by GRIPa_4 @1;
GRIP4 by GRIPb_4 (11);
GRIP4 by GRIPc_4(12) ;
[GRIPa_4] (34);
[GRIPb_4] (35);
[GRIPc_4](36) ;
cog4 by cfra4 @1;
cog4 by cfrb4 (6);
cog4 by cff_c4 (7);
cog4 by cfmd4 (8);
cog4 by cfre4 (9);
[cfra4] (27);
[cfrb4] (28);
[cff_c4] (29);
 [cfmd4] (30);
[cfre4] (31);
steady4 by Steada4 @1;
steady4 by Steadb4 (5);
[Steada4];
[Steadb4];
!Time 5
eye5 by Eyea5 @1;
eye5 by Eyeb5 (2);
eye5 by Eyec5 (3);
[Eyea5] (20);
[Eyeb5] (21);
[Eyec5] (22);
ear5 by Eara5 @1;
ear5 by Earb5 (4);
[Eara5] (23);
[Earb5] (24);
```

Appendix XIV (continued)
M-plus input file for model of balance using measures from ELSA

```
Gait5 by Gait_a5c @1;
Gait5 by Gait_b5c (10);
[Gait_a5c];
[Gait_b5c];
cog5 by cfra5 @1;
cog5 by cfrb5 (6);
cog5 by cff_c5 (7);
cog5 by cfmd5 (8);
cog5 by cfre5 (9);
[cfra5] (27);
[cfrb5] (28);
[cff_c5] (29);
[cfmd5] (30);
[cfre5] (31);
!Time 6
eye6 by Eyea6 @1;
eye6 by Eyeb6 (2);
eye6 by Eyec6 (3);
[Eyea6] (20);
[Eyeb6] (21);
[Eyec6] (22);
ear6 by Eara6 @1;
ear6 by Earb6 (4);
[Eara6] (23);
[Earb6] (24);
Gait6 by Gait_a6c @1;
Gait6 by Gait_b6c (10);
[Gait_a6c];
[Gait_b6c];
Chair6 by m6c @1;
m6c@0;
[m6c];
GRIP6 by GRIPa_6 @1;
GRIP6 by GRIPb_6 (11);
GRIP6 by GRIPc_6 (12);
[GRIPa_6];
[GRIPb_6];
[GRIPc_6];
```

Appendix XIV (continued)
M-plus input file for model of balance using measures from ELSA

```
cog6 by cfra6 @1;
cog6 by cfrb6 (6);
cog6 by cfre6(9);
[cfra6] (27);
[cfrb6] (28);
[cfre6] (31);
!time 1 correlations
CFRE1    WITH CFRB1;
!time 2 correlations
EYE2 WITH EYE;
EAR2 WITH EAR;
STEADY2 WITH STEADY;
COG2 WITH COG;
CFF_C2 WITH CFF_C1;
EYEB2 WITH EYEB1;
EARA2 WITH EARA1;
EARB2 WITH EARB1;
STEADB2 WITH STEADB1;
CFRA2 WITH CFRA1;
CFRE2 WITH CFRE1;
CFRE2 WITH CFRB2;
! time 3 correlations
EYE3 WITH EYE;
EYE3 WITH EYE2;
EAR3 WITH EAR;
EAR3 WITH EAR2;
GAIT3 WITH GAIT2;
EYEB3 WITH EYEB2;
EYEC3 WITH EYEC2;
EARA3 WITH EARA2;
EARB3 WITH EARB1;
EARB3 WITH EARB2;
! time 4 correlations
EYE4 WITH EYE3;
EAR4 WITH EAR3;
GAIT4 WITH GAIT3;
CHAIR4 WITH CHAIR2;
GRIP4 WITH GRIP2;
```

Appendix XIV (continued)
M-plus input file for model of balance using measures from ELSA

```
COG4 WITH COG;
COG4 WITH COG2;
STEADY4 WITH STEADY2;
EYEC4 WITH EYEC3;
EARA4 WITH EARA3;
EARB4 WITH EARB3;
STEADB4 WITH STEADB2;
CFRE4 WITH CFRB4;
CFF_C4 WITH CFF_C1;
CFF_C4 WITH CFF_C2;
EYE4 WITH EYE;
EYE4 WITH EYE2;
EAR4 WITH EAR;
EAR4 WITH EAR2;
STEADY4 WITH STEADY;
EARA4 WITH EARA1;
EARA4 WITH EARA2;
EARB4 WITH EARB2;
STEADA4 WITH STEADA1;
STEADB4 WITH STEADB1;
!TIME 5 CORRELATIONS
EYE5 WITH EYE4;
EAR5 WITH EAR4;
GAIT5 WITH GAIT4;
COG5 WITH COG4;
EYEA5 WITH EYEA4;
EARB5 WITH EARB4;
CFRA5 WITH CFRA4;
CFMD5 WITH CFMD4;
CFRE5 WITH CFRE4;
CFRE5 WITH CFRB5;
CFF_C5 WITH CFF_C4;
EYE5 WITH EYE3;
EAR5 WITH EAR3;
GAIT5 WITH GAIT3;
COG5 WITH COG;
COG5 WITH COG2;
EARB4 WITH EARB1;
```

Appendix XIV (continued)
M-plus input file for model of balance using measures from ELSA

```
EYEB5 WITH EYEB4;
EARA5 WITH EARA3;
EARB5 WITH EARB3;
CFF_C5 WITH CFF_C1;
CFF_C5 WITH CFF_C2;
! time 6 correlations
EYE6 WITH EYE5;
EAR6 WITH EAR2;
EAR6 WITH EAR4;
EAR6 WITH EAR5;
GAIT6 WITH GAIT5;
CHAIR6 WITH CHAIR4;
GRIP6 WITH GRIP2;
GRIP6 WITH GRIP4;
COG6 WITH COG5;
EYEA6 WITH EYEA5;
EYEB6 WITH EYEB5;
EYEC6 WITH EYEC5;
EARA6 WITH EARA5;
EARB6 WITH EARB4;
EARB6 WITH EARB5;
CFRA6 WITH CFRA5;
CFRE6 WITH CFRE5;
COG6 WITH COG4;
GAIT6 WITH GAIT4;
EAR6 WITH EAR3;
EYE5 WITH EYE2;
EAR5 WITH EAR2;
EYE6 WITH EYE4;
EAR6 WITH EAR;
CHAIR6 WITH CHAIR2;
COG6 WITH COG2;
EAR5 WITH EAR;
EYE6 WITH EYE3;
COG6 WITH COG;
!CORRELATED RESIDUAL BWTWEEN EAR & EYE-METHOD EFFECT
EAR WITH EYE;
EAR2 WITH EYE2;
EAR3 WITH EYE3;
EAR4 WITH EYE4;
```

Appendix XIV (continued)
M-plus input file for model of balance using measures from ELSA

```
EAR5 WITH EYE5;
EAR6 WITH EYE6;
balance by eye ear steady cog;
balance2 by eye2 ear2 steady2 cog2 GRIP2 Chair2 Gait2;
balance3 by eye3 ear3 Gait3;
balance4 by eye4 ear4 steady4 cog4 GRIP4 Chair4 Gait4;
balance5 by eye5 ear5 cog5 Gait5;
balance6 by eye6 ear6 cog6 GRIP6 Chair6 Gait6;
[balance];
[balance2];
[balance3];
[balance4];
[balance5];
[balance6];
output: modindices (all) stand;
```

Appendix XV
M-plus input for final model of physical activity and gait speed using ELSA

```
TITLE: ELSA - Final model for physical activity and Gait (gaitbc)
Data: file is elsa w1 and 2 3 4 5 6 merged_8 (9).dat;
define: age_1 =AGE1/10;
DEFINE: CENTER AGE_1 (GRANDMEAN);
VARIABLE: NAMES ARE idauniq etc.
USEVARIABLES ARE Gait_b2c Gait_b3c Gait_b4c Gait_b5c Gait_b6c AGE_1
SEX ADLa1 Falla1 Paina1 Alca2 EDU1 class3;
missing are all (-999, -9, -8, -1);
Grouping is class3 (1 = inactive 2 = low 3 = mod_vig);
analysis:ESTIMATOR = MLR;
iterations = 10000;
MODEL:
i s|Gait_b2c@0 Gait_b3c@1 Gait_b4c@2 Gait_b5c@3 Gait_b6c@4;
i on AGE_1 SEX ;!ADLa1 Falla1 Paina1 Alca2 EDU1;
s on AGE_1 SEX ;!ADLa1 Falla1 Paina1 Alca2 EDU1;
!Gait_b2c;
!Gait_b3c;
!Gait_b4c;
!Gait_b5c;
!Gait_b6c;
![Gait_b2c](10);
![Gait_b3c](11);
![Gait_b4c](12);
![Gait_b5c](13);
![Gait_b6c](14);
MODEL INACTIVE:
i s|Gait_b2c@0 Gait_b3c@1 Gait_b4c@2 Gait_b5c@3 Gait_b6c@4;
i on AGE_1 SEX ;!ADLa1 Falla1 Paina1 Alca2 EDU1;
s on AGE_1 SEX;! ADLa1 Falla1 Paina1 Alca2 EDU1;
!Gait_b2c ;!(1);
!Gait_b3c ;!(1);
!Gait_b4c ;!(1);
!Gait_b5c ;
!Gait_b6c ;
![Gait_b2c](10);
![Gait_b3c](11);
![Gait_b4c](12);
![Gait_b5c](13);
![Gait_b6c](14);
MODEL LOW:
i s|Gait_b2c@0 Gait_b3c@1 Gait_b4c@2 Gait_b5c@3 Gait_b6c@4;
i on AGE_1 SEX; ! ADLa1 Falla1 Paina1 Alca2 EDU1;
```

Appendix XV (continued)
M-plus input for final model of physical activity and gait speed using ELSA

```
s on AGE_1 SEX ;!ADLa1 Falla1 Paina1 Alca2 EDU1;
!Gait_b2c ;!(2);
!Gait_b3c ;!(2);
!Gait_b4c ;!(2);
!Gait_b5c ;
!Gait_b6c ;
[Gait_b2c];
[Gait_b3c];
![Gait_b4c](12);
![Gait_b5c](13);
![Gait_b6c](14);
MODEL MOD_VIG:
i s|Gait_b2c@0 Gait_b3c@1 Gait_b4c@2 Gait_b5c@3 Gait_b6c@4;
i on AGE_1 SEX ;!ADLa1 Falla1 Paina1 Alca2 EDU1;
s on AGE_1 SEX ;!ADLa1 Falla1 Paina1 Alca2 EDU1;
!Gait_b2c ;!(3);
!Gait_b3c ;!(3);
!Gait_b4c;! (3);
!Gait_b5c ;
!Gait_b6c ;
[Gait_b2c];
[Gait_b3c];
![Gait_b4c](12);
![Gait_b5c](13);
![Gait_b6c](14);
plot:
Type is plot1 plot2 plot3 ;
Series is Gait_b2c (1) Gait_b3c (2) Gait_b4c (3) Gait_b5c (4) Gait_b6c (5);
output: standardized modindices (ALL) SAMPSTAT;
```

Appendix XVI
M-plus input for final model of physical activity and steadiness using ELSA

```
TITLE: ELSA - Final model for physical activity and Steady (steada)
Data: file is elsa w1 and 2 3 4 5 6 merged_8 (9).dat;
define: AGE_1 = AGE1/10;
define: center AGE_1(grandmean);
VARIABLE: NAMES ARE idauniq ….
USEVARIABLES ARE
Steada1 Steada2 Steada4
AGE_1 SEX falla1 adla1 paina1 alca2 edu1 class3; !sleepa3
missing are all (-999, -9, -8, -1);
Grouping is class3 (1 = inactive 2 = low 3 = mod_vig);
ANALYSIS: estimator = mlr;
analysis: iterations = 10000;
Model: i s|Steada1@0 Steada2@1 Steada4*;
i on AGE_1 SEX falla1 adla1 paina1 alca2 edu1;
s on AGE_1 SEX falla1 adla1 paina1 alca2 edu1;
Model inactive:
i s|Steada1@0 Steada2@1 Steada4@2;
i on AGE_1 SEX falla1 adla1 paina1 alca2 edu1;
s@0;
Model low:
i s|Steada1@0 Steada2@1 Steada4@2;
i on AGE_1 SEX falla1 adla1 paina1 alca2 edu1;
s@0;
Model mod_vig:
i s|Steada1@0 Steada2@1 Steada4@2;
i on AGE_1 SEX falla1 adla1 paina1 alca2 edu1;
s@0;
savedata: file LGM_centering.sav;
plot: Type is plot1 plot2 plot3 ;
Series is Steada1(1) Steada2(2) Steada4(3);
output: standardized modindices (all) sampstat;
```

Appendix XVII
M-plus input for final model of physical activity and cognitive function (Cog) using ELSA

```
TITLE: ELSA - Final model for physical activity and Cognitive function
Data: file is elsa w1 and 2 3 4 5 6 merged_8 (9).dat;
DEFINE: AGE_1 = AGE1/10;
DEFINE: CENTER AGE_1 (GRANDMEAN);
VARIABLE: NAMES ARE idauniq
USEVARIABLES ARE cffc1 cffc2 cffc4 cffc5 AGE_1 SEX ADLa1 Falla1 Paina1
Alca2 EDU1 class3; missing are all (-999, -9, -8, -1);
Grouping is class3 (1 = inactive 2 = low 3 = mod_vig);
ANALYSIS: ESTIMATOR=MLR; ITERATIONS = 10000;
MODEL:i s|cffc1@0 cffc2@1 cffc4@2 cffc5@3;
i on AGE_1 SEX ADLa1 Falla1 Paina1 Alca2 EDU1;
s on AGE_1 SEX ADLa1 Falla1 Paina1 Alca2 EDU1;
cffc1;
cffc2;
cffc5;
MODEL INACTIVE:
i s|cffc1@0 cffc2@1 cffc4@2 cffc5@3;
i on AGE_1 SEX ADLa1 Falla1 Paina1 Alca2 EDU1;
cffc1(1);
cffc2(1);
cffc4(1);
cffc5(1);
S@0;
MODEL LOW:
i s|cffc1@0 cffc2@1 cffc4@2 cffc5@3;
i on AGE_1 SEX ADLa1 Falla1 Paina1 Alca2 EDU1;
s on AGE_1 SEX ADLa1 Falla1 Paina1 Alca2 EDU1;
cffc1(2);
cffc2(2);
cffc4(2);
cffc5(2);
MODEL MOD_VIG:
i s|cffc1@0 cffc2@1 cffc4@2 cffc5@3;
i on AGE_1 SEX ADLa1 Falla1 Paina1 Alca2 EDU1;
s on AGE_1 SEX ADLa1 Falla1 Paina1 Alca2 EDU1;
cffc1(3);
cffc2(3);
cffc4(3);
cffc5(3);
plot: Type is plot1 plot2 plot3; Series is cffc1 (1) cffc2 (2) cffc4 (3) cffc5 (4);
output: standardized modindices (ALL) SAMPSTAT;
```

Appendix XVIII
Detailed comparison of the measures used in TILDA, ELSA and NICOLA longitudinal studies

Body system	Measures	TILDA	ELSA	NICOLA
Sensory measures	Vision	One question: *Is your eyesight (using glasses or corrective lenses) ... (1=excellent; 2=very good; 3=good; 4=fair; 5=poor)	Three questions: *1) How is your eyesight? (using glasses or corrective lenses) ... 2) How is your eyesight seeing at a distance? 3) How is your eyesight seeing close up? (1=excellent; 2=very good; 3=good; 4=fair; 5=poor)	Two questions: *1) At the present time is your eyesight? 2) How much difficulty do you have doing hobbies that require you to see close up? (1=excellent; 2=very good; 3=good; 4=fair; 5=poor)
	Hearing	One question: Is your hearing (with or without a hearing aid) (1==excellent; 2=very good; 3=good; 4=fair; 5=poor)	Two questions: 1) How is your hearing? (1=excellent; 2=very good; 3=good; 4=fair; 5=poor) 2) Do you find it difficult to follow conversation with background noise? (Yes/no)	n/a
Cognitive measure	Executive function	One measure Mini mental state exam, a 30 item scale	Five questions 1) Date recall 2) Immediate word recall 3) No. animals recalled 4) Remembering to write initials 5) Delayed word recall Score	Two questions 1) Immediate word recall 2) No. animals recalled

Appendix XVIII (continued)
Detailed comparison of the measures used in TILDA, ELSA and NICOLA longitudinal studies

Balance measures	Body system	TILDA	ELSA	NICOLA
Neuro-muscular measures	Strength	Handgrip test Highest score dominant hand (kg) One averaged measure	Handgrip test Highest score dominant hand (kg) Three measures	Handgrip test Highest score dominant hand (kg) Two measures
	Strength, mobility, gait speed	Timed Up & Go (secs) 1 measure	Two measures 1)Chair rise Test (sec) highest - one measure 2)Gait speed (3km) (sec) – two scores	Timed Up & Go (secs) 1 measure
Physical activity	N/A	Three questions: (IPAQ) During last 7 days how many mins did you do vigorous/moderate/walk? Minutes per week	Three questions regarding free-living activity: How often do you take part in sports or activities that are vigorous/moderately energetic/mildly energetic? (1=more than once per week; 2=once per week; 3=one to three times per week; 4=hardly ever)	Three questions: (IPAQ) During last 7 days how many mins did you do vigorous/moderate/walk? Minutes per week

Appendix XIX
M-plus input for comparison of measures of grip, eye, and steady from TILDA, ELSA, and NICOLA studies

```
TITLE: multi-group analysis for TILDA, ELSA and NICOLA
Data: file is combined_3.dat;
variable: names are
idauniq Eyea1 Eyeb1 Eyec1 Eara1 Steada1 Steadb1 Stead3w1Gaita1 Gaitb1
cfra1 cfrb1 cffc1 cfmd1 cfre1 MMSE dhager1 Gait_a1 Gait_b1 aGait1 bGait1
ADLa1 Falla1
Paina1 EDU1 PAa1 PAb1 PAc1 Agec_1 Agec_2 Agec_3 SEX MED class3
GRIPa2 GRIPb2 GRIPc2 TUGR PAMET m2c COUNTRY;
USEVARIABLES ARE
Eyea1 Steada1 GRIPa2 GRIPb2 Country;
missing are all (-999, -9, -8, -1);
Grouping is country (2 = elsa 1 = tilda 3 = NICOLA);
ANALYSIS: estimator = mlr;
MODEL:
Eye by Eyea1@1;
Eyea1@0;
Steady by Steada1@1;
Steada1@0; (6);
GRIP by GRIPa2@1 (7);
GRIP by GRIPb2@1 (7);
[GRIPa2] (8);
[GRIPb2];
GRIPa2;
GRIPb2;
Model elsa:
Steady with eye;
GRIP with eye;
Grip with Steady (12);
Model tilda:
Steady with eye;
GRIP with eye;
Grip with Steady;
Model NICOLA:
Steady with eye;
GRIP with eye;
Grip with Steady (12);
output: standardized modindices;
```

Appendix XX
M-plus input for comparison of measures of Grip, Eye, and Steady from TILDA, ELSA, and NICOLA studies (correlations)

```
TITLE: multi-group correlations TILDA, ELSA, NICOLA
Data: file is combined_3.dat;
variable: names are
idauniq Eyea1 Eyeb1 Eyec1 Eara1 Steada1 Steadb1 Stead3w1 Gaita1 Gaitb1
cfra1 cfrb1 cffc1 cfmd1 cfre1 MMSE dhager1 Gait_a1 Gait_b1 aGait1 bGait1
ADLa1 Falla1
Paina1 EDU1 PAa1 PAb1 PAc1 Agec_1 Agec_2 Agec_3 SEX MED class3
GRIPa2 GRIPb2 GRIPc2 TUGR PAMET m2c COUNTRY;
USEVARIABLES ARE
Eyea1 Eyec1Steada1 GRIPa2 cfrb1 cffc1;
missing are all (-999, -9, -8, -1);
USEOBSERVATION ARE (country EQ 3);
ANALYSIS: estimator = mlr;
MODEL:
Eyea1 with Eyec1;
Eyea1 with Steada1;
Eyec1 with Steada1;
Steada1 with GRIPa2;
Eyea1 with GRIPa2;
Eyea1 with cfrb1;
Eyec1 with cfrb1;
GRIPa2 with cffc1;
cfrb1 with cffc1;
cfrb1 with GRIPa2;
cffc1 with Eyea1;
cffc1 with Eyec1;
cffc1 with Steada1;
Steada1 with cfrb1;
GRIPa2 with cfrb1;
Eyec1 with GRIPa2;
output: standardized modindices;
```