Effectiveness of a Dementia-Specific Multimodal Exercise Program on Motor and Gait Performance in Individuals with Dementia – a Randomized Controlled Trial

Zur Erlangung des akademischen Grades einer

DOKTORIN DER PHILOSOPHIE (Dr. phil.)

von der KIT-Fakultät für Geistes- und Sozialwissenschaften des

Karlsruher Instituts für Technologie (KIT)

angenommene

DISSERTATION

von

Sandra Trautwein

KIT-Dekan: Prof. Dr. Michael Schefczyk

1. Gutachter: Prof. Dr. Alexander Woll

2. Gutachter: Prof. Dr. Thorsten Stein

Tag der mündlichen Prüfung: 18. Dezember 2019

Preface

This thesis comprises the following manuscripts, published or under review in peer-reviewed journals:

Manuscript I

Trautwein, S., Barisch-Fritz, B., Scharpf, A., Bossers, W., Meinzer, M., Steib, S., Stein, T., Bös, K., Stahn, A., Niessner, C., Altmann, S., Wittelsberger, R., & Woll, A. (2019). Recommendations for assessing motor performance in individuals with dementia: Suggestions of an expert panel - a qualitative approach. *European Review of Aging and Physical Activity*, *16*(5). https://doi.org/10.1186/s11556-019-0212-7

Manuscript II

Trautwein, S., Maurus, P., Barisch-Fritz, B., Hadzic, A., & Woll, A. (2019). Recommended motor assessments based on psychometric properties in individuals with dementia: A systematic review. *European Review of Aging and Physical Activity*, *16*(20). https://doi.org/10.1186/s11556-019-0228-z

Manuscript III

Trautwein, S., Scharpf, A., Barisch-Fritz, B., Niermann, C., & Woll, A. (2017). Effectiveness of a 16-week multimodal exercise program on individuals with dementia: Study protocol for a multicenter randomized controlled trial. *JMIR Research Protocols*, *6*(3), e35. https://doi.org/10.2196/resprot.6792

Manuscript IV

Barisch-Fritz, B., Trautwein, S., Scharpf, A., Krell-Roesch, J., & Woll, A. (submitted). Effects of a 16-week multimodal exercise program on motor performance in individuals with dementia: A multicenter randomized controlled trial.

Manuscript V

Trautwein, S., Barisch-Fritz, B., Scharpf, A., Ringhof, S., Stein, T., Krell-Roesch, J., & Woll, A. (submitted). Effects of a 16-week multimodal exercise program on gait performance in individuals with dementia: A multicenter randomized controlled trial.

Abstract

Background: Dementia is not only characterized by cognitive changes but also has an impact on motor performance. Compared to cognitively unimpaired older adults, individuals with dementia (IWD) show a decreased motor performance and are frequently affected by motor impairments, e.g. in gait, balance, mobility, and strength. Motor impairments are related to an increasing loss of independence and a high need for care. Furthermore, they reduce quality of life and represent huge challenges for affected individuals and their caregivers. Considering the high prevalence of dementia, there is a great need for effective treatments of this disease. So far, there are no causal therapy strategies, and pharmacological approaches are associated with various side effects. Therefore, non-pharmacological strategies such as physical activity are becoming increasingly important. Numerous studies investigate the effectiveness of physical activity on motor, gait, and cognitive performance in IWD. Due to methodological limitations, however, evidence is still limited. This thesis focuses on the investigation of the effectiveness of physical activity on motor and gait performance in IWD. It pursues the following objectives: 1) to establish a high-quality methodological approach to investigate the effectiveness of physical activity in IWD, and 2) to perform a high-quality randomized controlled trial examining the effects of a dementia-specific multimodal exercise program on motor and gait performance in IWD.

Methods: The above-mentioned objectives were carried out on the basis of two primary and four secondary research questions, from which seven hypotheses were derived. In a first step, a high-quality methodological approach was established based on a comprehensive examination of the current state of research and previous studies. It focused on the evaluation of the adequateness of motor assessments previously applied in IWD. In addition, a study design, which fulfills quality criteria derived from the current state of research and considers specific characteristics of IWD, was developed within the framework of a study protocol. In a second step, this high-quality methodological approach was applied in a randomized controlled trial. For this purpose, 319 individuals with mild to moderate dementia were randomly assigned to an intervention or control group. The intervention group participated in a 16-week multimodal exercise program specifically tailored to IWD. In order to examine the effects of this multimodal exercise program, motor and gait performance of all participants was assessed with a comprehensive assessment battery before and after the intervention. In addition to

determining time*group effects on motor and gait performance, responder-non-responder-analyses and multiple linear regression models were used to identify prerequisites and impacts of changes in underlying motor and cognitive performance on changes in gait performance.

Results: Based on the limitations of previous studies as identified in recent reviews, various methodological quality criteria were defined. In addition to a general high-quality methodological approach, the use of adequate motor assessments and interventions tailored to IWD were determined as key issues. A qualitative examination of motor assessment used in previous studies emphasized the importance of adequate assessments for the target group. In addition, the use of a sequential approach, a selection of eight motor assessments, and the allowance of standardized repetitions of instructions as well as use of walking aids were recommended. Based on a quantitative evaluation, eight motor assessments, which are characterized by sufficient relative reliability, assumed sensitivity, and frequent use in previous studies were recommended. Considering insufficient absolute reliability, the adequateness of these motor assessments, however, is limited for examining intra-individual changes. Taking into account these findings on adequate motor assessments, a high-quality study design was developed. In addition, a randomized controlled trial examining the effectiveness of a multimodal exercise program showed no statistically significant time* group effects on motor and gait performance in IWD. In further exploratory analyses, differences in baseline performance of gait, mobility, strength, and severity of cognitive impairments were observed between positive, non-, and negative responders. Moreover, the impacts of changes in strength and function of lower limbs, mobility, executive function, attention, and working memory explained up to 39.4 % of the variance of changes in gait performance.

Conclusion: This thesis presents an important contribution to improving the methodological quality of studies investigating the effectiveness of physical activity on motor and gait performance in IWD. It includes a comprehensive examination of adequate motor assessments and suggests an overall high-quality methodological approach, which can be useful for future studies. Nevertheless, additional investigations are required to further improve methodological approaches in this field of research. For example, the theoretically established recommendations on motor assessments must be evaluated in appropriate studies and new motor assessments specifically tailored to IWD are needed. In addition to these methodological recommendations and research

perspectives, this thesis provides important findings for designing and implementing future physical activity interventions that consider the specific characteristics of the target group. The hypothesis that a dementia-specific multimodal exercise program has a positive effect on motor and gait performance in IWD could not be confirmed. This shows that a standardized physical activity intervention is only suitable to a limited extent for this heterogeneous target group. A proportion of up to 40 % of positive responders as well as the identification of necessary prerequisites and changes in underlying motor and cognitive performance allow the assumption that the effectiveness of physical activity is dependent on more complex mechanisms and requires the consideration of individual characteristics. Therefore, approaches from individualized medicine offer promising perspectives. In line with this, it is assumed that the effectiveness of physical activity can be increased if appropriate interventions are tailored to individual prerequisites of participants. Such approaches must be specified and investigated in future studies. For example, the identification of specific characteristics represents a promising research perspective that allows to better describe IWD and to define different clusters of participants with various prerequisites. In addition to the derivation of numerous research perspectives, the results of this thesis also have important application-oriented implications that emphasize its practical relevance. These include indications on designing and implementing physical activity interventions for IWD, but also the concrete application of recommendations on motor assessments and the multimodal exercise program in everyday life of care facilities. In summary, physical activity offers great potential for dealing with dementia in our society. High-quality research such as the manuscripts included in this thesis contribute to exploit this potential and may enable IWD and their caregivers to benefit from the positive effects of physical activity.

Zusammenfassung

Hintergrund: Demenzerkrankungen kennzeichnen sich nicht nur durch Veränderungen der Kognition, sondern beeinflussen auch die motorische Leistungsfähigkeit. Im Vergleich zu kognitiv gesunden älteren Menschen, zeigen Personen mit Demenz eine reduzierte motorische Leistung und sind häufiger von motorischen Beeinträchtigungen betroffen, z. B. im Bereich des Ganges, des Gleichgewichts, der Mobilität oder der Kraft. Motorische Beeinträchtigungen führen u. a. zu einem zunehmenden Verlust der Selbstständigkeit und einem erhöhten Pflegebedarf. Sie haben einen negativen Einfluss auf die Lebensqualität und stellen eine große Herausforderung für Betroffene und ihre Angehörigen dar. Unter Berücksichtigung der hohen Demenzprävalenz besteht ein großer Bedarf an effektiven Behandlungsmöglichkeiten von Demenzerkrankungen. Bislang gibt es keine ursächlichen Therapiestrategien und medikamentöse Maßnahmen gehen mit verschiedenen Nebenwirkungen einher. Daher gewinnen nichtmedikamentöse Ansätze wie z. B. körperliche Aktivität zunehmend an Bedeutung. Zahlreiche Studien untersuchen den Einfluss körperlicher Aktivität auf die Motorik, den Gang und die Kognition bei Personen mit Demenz. Aufgrund methodischer Limitationen dieser Studien liegt bisher jedoch keine gesicherte Evidenz vor. Im Mittelpunkt dieser Thesis steht die Untersuchung der Effekte körperlicher Aktivität auf die Motorik und den Gang bei Personen mit Demenz. Hierbei werden zwei Ziele verfolgt: 1) die Erarbeitung eines qualitativ hochwertigen methodischen Ansatzes zur Untersuchung der Effektivität körperlicher Aktivität bei Personen mit Demenz und 2) die Durchführung einer qualitativ hochwertigen randomisierten kontrollierten Studie, die die Effekte eines demenzspezifischen multimodalen Bewegungsprogrammes auf die Motorik und den Gang bei Personen mit Demenz überprüft.

Methoden: Die Bearbeitung der oben genannten Ziele erfolgte anhand von zwei übergeordneten und vier untergeordneten Fragestellungen, aus denen sieben Hypothesen abgeleitet wurden. In einem ersten Schritt wurde aufbauend auf einer umfangreichen Analyse des aktuellen Forschungstandes sowie bisherigen Studien ein qualitativ hochwertiges methodisches Vorgehen erarbeitet. Der Schwerpunkt lag hierbei auf der Bewertung der Eignung bisher verwendeter motorischer Testverfahren für Personen mit Demenz. Darüber hinaus wurde im Rahmen eines Studienprotokolls ein Studiendesign entwickelt, das aus dem Stand der Forschung abgeleitete Qualitätskriterien erfüllt und

die Besonderheiten von Personen mit Demenz berücksichtigt. Dieses erarbeitete qualitativ hochwertige methodische Vorgehen wurde in einem zweiten Schritt im Rahmen einer randomisierten kontrollierten Studie eingesetzt. Hierfür wurden 319 Personen mit leichter bis mittelschwerer Demenz randomisiert einer Interventions- oder Kontrollgruppe zugeteilt. Die Interventionsgruppe absolvierte ein 16-wöchiges multimodales Bewegungsprogramm, das die Besonderheiten der Zielgruppe berücksichtigt. Zur Überprüfung der Effekte dieses Bewegungsprogrammes wurden die motorische und die Gangleistung aller Teilnehmenden vor und nach der Intervention mit einer umfangreichen Testbatterie überprüft. Neben der Bestimmung von Zeit*Gruppen Effekten auf die Motorik und den Gang, wurden mithilfe von Responder-Nicht-Responder-Analysen und multiplen linearen Regressionen Voraussetzungen und zugrundeliegende Anpassungen identifiziert, die mit den beobachteten Änderungen der Gangleistung zusammenhängen.

Ergebnisse: Anhand von Limitationen bisheriger Studien, die in aktuellen systematischen Reviews identifiziert wurden, wurden verschiedene methodische Qualitätskriterien definiert. Neben einem allgemeinen qualitativ hochwertigen methodischen Vorgehen, wurden v. a. die Verwendung von geeigneten motorischen Testverfahren und an die Zielgruppe angepassten Interventionen als wesentliche Qualitätsmerkmale ermittelt. Eine qualitative Überprüfung bisher eingesetzter motorischer Tests unterstreicht die Bedeutung von für die Zielgruppe geeigneten Verfahren. Zudem wurde die Verwendung eines sequenziellen Ansatzes, eine Auswahl an acht motorischen Testverfahren sowie das Zulassen von standardisierten Wiederholungen der Instruktionen und des Einsatzes von gängigen Hilfsmitteln empfohlen. Basierend auf einer quantitativen Bewertung wurden acht motorische Tests empfohlen, die sich durch ausreichende relative Reliabilität, angenommene Sensitivität und häufige Verwendung in bisherigen Studien auszeichnen. Allerdings sind diese Testverfahren aufgrund ihrer ungenügenden absoluten Reliabilität nur begrenzt geeignet um intraindividuelle Veränderungen der motorischen Leistung zu erfassen. Unter Berücksichtigung dieser Erkenntnisse zu geeigneten motorischen Testverfahren, wurde ein qualitativ hochwertiges Studiendesign entwickelt. Darüber hinaus zeigte eine randomisierte kontrollierte Studie zur Überprüfung der Wirksamkeit körperlicher Aktivität keine statistisch signifikanten Zeit*Gruppen Effekte auf die motorische und die Gangleistung von Personen mit Demenz. In weiteren explorativen Analysen wurden Unterschiede in der Ausgangsleistung des

Ganges, der Mobilität, der Kraft und dem Grad der kognitiven Beeinträchtigung zwischen Positiv-, Nicht- und Negativ-Respondern beobachtet. Zudem erklären Änderungen der Kraft und Funktion der unteren Extremitäten, der Mobilität, der exekutiven Funktion, der Aufmerksamkeit und des Arbeitsgedächtnisses 39.4 % der Varianz der Änderung der Gangleistung.

Schlussfolgerungen: Die vorliegende Thesis leistet einen wichtigen Beitrag zur Verbesserung der methodischen Qualität von Studien zur Überprüfung der Wirksamkeit körperlicher Aktivität auf die motorische und die Gangleistung von Personen mit Demenz. Sie beinhaltet eine umfangreiche Auseinandersetzung mit motorischen Testverfahren, die für Personen mit Demenz geeignet sind, und schlägt ein qualitativ hochwertiges allgemeines methodisches Vorgehen vor, an dem sich zukünftige Studien orientieren können. Dennoch besteht weiterer Forschungsbedarf zur Verbesserung des methodischen Vorgehens im Forschungsfeld. Beispielsweise müssen die theoretisch erarbeiteten Empfehlungen für motorische Testverfahren in entsprechenden Studien überprüft und neue motorische Tests speziell für Personen mit Demenz entwickelt und untersucht werden. Neben diesen methodischen Empfehlungen und Forschungsperspektiven liefert die Thesis wichtige Erkenntnisse für die Gestaltung und Umsetzung von Bewegungsprogrammen, die an die Besonderheiten der Zielgruppe angepasst sind. Die Hypothese, dass ein demenzspezifisches multimodales Bewegungsprogramm sich positiv auf die motorische und die Gangleistung von Personen mit Demenz auswirkt, konnte nicht bestätigt werden. Damit lässt sich zeigen, dass ein standardisiertes Bewegungsprogramm für die heterogene Zielgruppe von Personen mit Demenz nur begrenzt geeignet ist. Ein Anteil von bis zu 40 % an Positiv-Respondern sowie die Identifizierung von notwendigen Voraussetzungen und Anpassungen der zugrundeliegenden motorischen und kognitiven Leistung, erlauben die Annahme, dass die Wirksamkeit körperlicher Aktivität auf komplexeren Mechanismen beruht und eine Berücksichtigung des Individuums erfordert. Daher bieten Ansätze aus der individualisierten Medizin vielversprechende Perspektiven. Es wird davon ausgegangen, dass die Effektivität körperlicher Aktivität gesteigert werden kann, wenn Interventionen an die individuellen Voraussetzungen der Teilnehmenden angepasst sind. Entsprechende Ansätze müssen in zukünftigen Studien konkretisiert und untersucht werden. Beispielsweise stellt die Identifikation charakteristischer Merkmale eine vielversprechende Forschungsperspektive dar, die es erlaubt Personen mit Demenz besser zu

beschreiben und verschiedene Cluster mit unterschiedlichen Anforderungen zu definieren. Neben der Ableitung zahlreicher Forschungsperspektiven erlauben die Ergebnisse dieser Thesis wichtige anwendungsbezogene Implikationen, die ihre praktische Relevanz unterstreichen. Hierzu zählen grundlegende Hinweise zur Gestaltung und Umsetzung von Bewegungsprogrammen für Personen mit Demenz, aber auch die konkrete Anwendung der Empfehlungen zu motorischen Testverfahren und des multimodalen Bewegungsprogrammes im Alltag von Pflegeeinrichtungen. Körperliche Aktivität bietet ein großes Potential für den Umgang mit Demenzerkrankungen in unserer Gesellschaft. Qualitativ hochwertige Forschung wie diese Thesis tragen dazu bei, dass dieses Potential ausgeschöpft werden kann und möglichst viele von der Erkrankung betroffene Personen und ihre Angehörigen die Möglichkeit haben von den positiven Effekten körperlicher Aktivität zu profitieren.

Contents

List of figures

List of tables

List of tables - appendix

List of abbreviations

6m WT:	6-meter/metre walk test
6min WT:	6-minute walk test
30s CST:	30-second chair stand test
AD:	Alzheimer's disease
ADL:	activities of daily living
BBS:	Berg Balance Scale
CONSORT:	Consolidated Standards of Reporting Trials
FICSIT-4:	Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4
FR:	Functional Reach Test
GMWT:	Groningen Meander Walking Test
ICD-10:	International Statistical Classification of Diseases and Related Health Problems, 10th Revision/Tenth Edition
IG:	intervention group
IWD:	individuals with dementia
MCI:	mild cognitive impairment
MEP:	multimodal exercise program
N/A:	not available
POMA:	Performance Oriented Mobility Assessment
PPT:	Physical Performance Test
RCT:	randomized/randomised controlled trial/s
SPPB:	Short Physical Performance Battery
STS:	Sit-to-Stand/sit-to-stand
TUG:	Timed Up & Go Test
VD:	vascular dementia
WT:	walk test/s

Further abbreviations additionally used within manuscripts

4m WT:	4-metre walk test
5x STS:	Five Times Sit-to-Stand Test
ACSID:	Assessment of Compensatory Sit-to-Stand Maneuvers in People With Dementia
ADAS-Cog:	Alzheimer Disease Assessment Scale–Cognitive Subscale
ANOVA:	analysis of variance
BMI:	Body Mass Index
CERAD:	Consortium to Establish a Registry for Alzheimer's Disease
CG:	control group
CIRS:	Cumulative Illness Rating Scale
COSMIN:	COnsensus-based Standards for the selection of health Measurement INstruments
E-ADL Test:	Erlangen Test of Activities of Daily Living
ICC:	intraclass correlation coefficient/s
KIT:	Karlsruhe Institute of Technology
$MDC_{95\%}$:	percentage minimal detectable change/s at 95 % confidence interval
MMSE:	Mini-Mental State Examination
SFT:	Senior Fitness Test
SPIRIT:	Standard Protocol Items: Recommendations for Interventional Trials

1 Introduction

1.1 Research problem and relevance

The worldwide trend of the aging population results in a dramatic increase of individuals with non-communicable diseases, such as dementia (Alzheimer's Disease International, 2015; Alzheimer's Association, 2019; Lin & Lewis, 2015). In 2018/2019, 50 million people worldwide were living with dementia. According to current estimates, this prevalence will further increase reaching 152 million cases in 2050 (Alzheimer's Disease International, 2018, 2019). "Dementia is the greatest global challenge for health and social care in the 21st century" (Livingston et al., 2017, p. 2673). It is associated with economic impacts, high demand for care, and burdens for affected families (Alzheimer's Disease International, 2015; Alzheimer's Association, 2019; Du et al., 2018; Livingston et al., 2017; Schulze, van den Bussche, Kaduszkiewicz, Koller, & Hoffmann, 2015). For example, the total estimated worldwide costs of dementia amounted to one trillion US Dollars in 2018, which are assumed to double until 2030 (Alzheimer's Disease International, 2018). Moreover, dementia represents one of the major causes of disability and dependency in older adults and thus considerably compromises the quality of life of individuals with dementia (IWD) (Alzheimer's Association, 2019; Hausdorff & Buchman, 2013; Livingston et al., 2017).

Dementia is primarily associated with cognitive impairments and behavioral changes (Waldemar et al., 2007). However, it also affects motor and functional performance, such as balance, mobility, gait, strength and function of lower limbs (Harlein et al., 2009, 2009; Ijmker & Lamoth, 2012; Pettersson, Olsson, & Wahlund, 2005; Suttanon et al., 2012). Based on the importance of motor and cognitive performance for safe and effective gait, this thesis considers gait as an example of functional performance frequently impaired in IWD (Beauchet et al., 2008). As motor and gait impairments impede autonomy in everyday life, they belong to the most problematic aspects of the disease for IWD and their caregivers (Hageman & Thomas, 2002). More detailed, motor and gait impairments are related to loss of independence, immobility, impaired performance of activities of daily living (ADL), and increased risk of falls. As a consequence, IWD are frequently affected by functional decline, disability, admission to emergency department, institutionalization, and mortality (Allali & Verghese, 2017; Amboni, Barone, & Hausdorff, 2013; Eggermont et al., 2010; Harlein et al., 2009; Mazoteras Muñoz et al., 2010; Nakayama, Suzuki, & Hamaguchi, 2019; Suttanon et al.,

2012; Thomas, Vandenberg, & Potter, 2002). Together with the above-mentioned general impacts of dementia, these consequences emphasize the need for effective strategies for reducing dementia-related motor and gait impairments (Brett, Traynor, & Stapley, 2016; Hausdorff & Buchman, 2013; McGough, Logsdon, Kelly, & Teri, 2013).

Currently, there is no cure for neurodegenerative and vascular dementia (VD) and available medication does not enable disease-modifying therapies. Dementia-specific drugs, e.g. cholinesterase inhibitors or memantine, as well as antidepressants or neuroleptics, are used for symptomatic treatment (Hugo & Ganguli, 2014; Versijpt, 2014). Besides, nonpharmacological treatments such as behavior management therapies are effective in treating behavioral and psychological symptoms (Dyer, Harrison, Laver, Whitehead, & Crotty, 2018). With respect to motor and gait impairments, physical activity is the preferred nonpharmacological treatment option due to its wide range of benefits, few side effects, and low economic burden (Du et al., 2018). In cognitively unimpaired older adults, the effectiveness of physical activity on motor and gait performance is supported by various studies (Chodzko-Zajko et al., 2009; Giné-Garriga, Roqué-Fíguls, Coll-Planas, Sitjà-Rabert, & Salvà, 2014; Hortobágyi et al., 2015; Latham, Bennett, Stretton, & Anderson, 2004; Liu & Latham, 2009; Liu, Shiroy, Jones, & Clark, 2014; Mian, Baltzopoulos, Minetti, & Narici, 2007). Even though a growing body of research has demonstrated a positive impact of physical activity on motor and gait performance in IWD, the evidence in this population is still limited, as previous studies frequently show methodological limitations and high risk of bias (Du et al., 2018; Farina, Rusted, & Tabet, 2014).

In addition to performing further studies, it is thus necessary to focus on methodological approaches and to develop high-quality study designs. Accordingly, this thesis aims to establish a high-quality methodological approach for investigating the effectiveness of physical activity on motor and gait performance in IWD and to perform a high-quality randomized controlled trial (RCT). Besides determining the overall effects of a dementia-specific multimodal exercise program (MEP), this RCT also considers characteristics of responders of the MEP and impacts of changes in underlying motor and cognitive performance on changes in gait performance. Therefore, this thesis provides a valuable contribution to improving research practices and developing effective physical activity interventions specifically tailored to IWD.

1.2 Structure of the thesis

The overarching aim of this thesis is to investigate the effectiveness of physical activity on motor and gait performance in IWD. Figure 1 summarizes its structure.

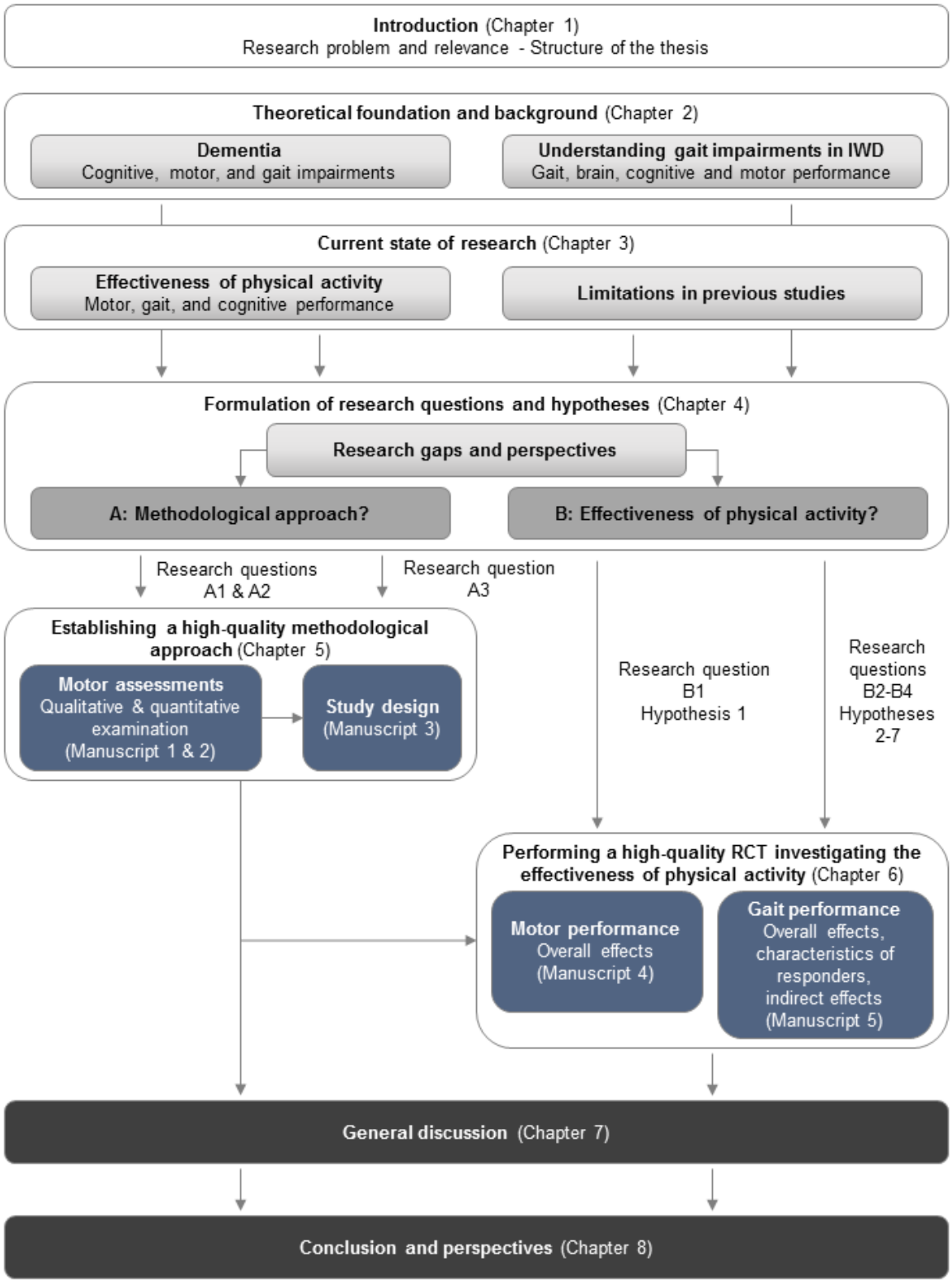


Figure 1. Structure of the thesis (IWD: individuals with dementia, RCT: randomized controlled trial).

Chapter 1 provides a brief introduction to the overall research problem and its relevance. The aim to investigate the effectiveness of physical activity on motor and gait performance in IWD is pursued from two perspectives: i) establishing a high-quality methodological approach and ii) performing a high-quality RCT. Pertinent research questions and hypotheses are based on theoretical considerations and the current state of research.

Chapter 2 builds a theoretical foundation and summarizes cognitive, motor, and gait impairments associated with dementia as well as influences on dementia-specific gait impairments and potential underlying mechanisms. Following, Chapter 3 gives a detailed overview of the current state of research.

With respect to the aim to establish a high-quality methodological approach, Chapters 2 and 3 intent to build a theoretical basis to answer the following questions:

- Which cognitive and motor impairments in IWD need to be considered when selecting motor assessments and developing physical activity interventions specifically tailored to IWD? (Chapters 2.1.1 and 2.1.2)
- Is there any relation between motor and cognitive performance in IWD, which supports the application of interventions combining physical and cognitive activity? (Chapter 2.2, using the example of gait performance)
- Which limitations of previous studies need to be considered when designing high-quality studies to investigate the effectiveness of physical activity on motor and gait performance in IWD? (Chapter 3.4)

The theoretical foundation and the current state of research also form an important basis for performing a high-quality RCT investigating the effectiveness of physical activity on motor and gait performance in IWD. This RCT firstly considers overall effectiveness on motor performance, and secondly focuses on gait performance aiming to determine overall effectiveness, characteristics of responders, and impacts of changes in underlying motor and cognitive performance on changes in gait performance. With respect to the aim to investigate the overall effectiveness on motor performance, Chapters 2 and 3 summarize findings which allow to answer the questions presented below:

- Which motor impairments that can potentially be influenced by physical activity occur in IWD? (Chapter 2.1.2)
- What is known about the effectiveness of physical activity on motor performance in IWD? (Chapter 3.1)

Focusing on the aim to determine overall effectiveness, characteristics of responders, and impacts of changes in underlying motor and cognitive performance on gait performance, requires a theoretical foundation with regard to the subsequent questions:

- In which spatiotemporal gait parameters, that are potentially sensitive to physical activity, impairments do occur in IWD? (Chapter 2.1.3)
- Are there any differences in cognitive, motor, and gait performance between individuals with different severities and etiologies of dementia, potentially influencing the effectiveness of physical activity on gait performance? (Chapter 2.1)
- Which cognitive and motor functions are associated with gait performance in IWD and thus may contribute to explain intervention-induced changes in gait performance? (Chapters 2.2.2 and 2.2.3)
- What is known about the effectiveness of physical activity on gait performance in IWD? (Chapter 3.2)
- Are there any indications for the effectiveness of physical activity on cognitive performance in IWD, which allow assuming intervention-induced cognitive impacts on changes in gait performance? (Chapter 3.3)

Chapter 4 focuses on the formulation of research questions and derived hypotheses. It starts with a summary of research gaps and perspectives related to the current state of research. Based on this summary two primary and four secondary research questions are established. These are investigated on the basis of seven hypotheses, which are built on theoretical foundations.

Elaborated research questions and derived hypotheses are examined in five research articles presented in Chapters 5 and 6. Chapter 5 focuses on establishing a high-quality methodological approach (primary research question A), while Chapter 6 reports findings of a high-quality RCT investigating the effectiveness of physical activity on motor and gait performance in IWD (primary research question B).

The subsequent general discussion in Chapter 7 aims to answer the research questions. It summarizes and critically discusses related findings of research articles. The thesis ends with conclusions in Chapter 8 highlighting the relevance of observed findings, possible research perspectives, and practical implications.

2 Theoretical foundation and background

2.1 Dementia

According to the International Statistical Classification of Diseases and Related Health Problems, 10th Revision (ICD-10) criteria dementia (F00-F03) is defined as

> […] syndrome due to disease of the brain, usually of a chronic or progressive nature, in which there is disturbance of multiple higher cortical functions, including memory, thinking, orientation, comprehension, calculation, learning capacity, language, and judgement. Consciousness is not clouded. The impairments of cognitive function are commonly accompanied, and occasionally preceded, by deterioration in emotional control, social behaviour, or motivation. This syndrome occurs in Alzheimer disease, in cerebrovascular disease, and in other conditions primarily or secondarily affecting the brain. (World Health Organization, 2016)[1]

The definition shows that dementia is not a single disease, but a syndrome including various etiologies. Related to different severities and etiologies of dementia, several initial predisposing factors and lifelong events, IWD are characterized by large heterogeneity with respect to cognitive, motor, and functional impairments (Cohen-Mansfield, 2000; Mitnitski, Graham, Mogilner, & Rockwood, 1999; Suttanon, Hill, Said, & Dodd, 2010).

Accounting for 50-60 % of dementia cases, Alzheimer's disease (AD) is the most frequent etiology of dementia followed by VD and mixed dementia (Brunnström, Gustafson, Passant, & Englund, 2009; Kumfor, Halliday, & Piguet, 2017; Livingston et al., 2017; Salardini, 2019). Other etiologies like Lewy body dementia or frontotemporal dementia are less common (Brunnström et al., 2009; Livingston et al., 2017; Salardini, 2019) and thus are not considered more detailed herein. The neuropathological hallmarks of AD include neuritic plaques of amyloid-beta protein and hyperphosphorylated tau neurofibrillary tangles. The progressive pathology of AD starts with aggregations of intraneuronal tau in limbic regions as well as the temporal lobe and continues with plaque depositions in associated cortices. Consequently, further and more widespread cortical tangle formations arise. It is suggested that these pathologic processes cause neurodegenerations resulting in increasing brain atrophy over time (Kumfor et al., 2017). However, it is not finally clarified if they actually induce neurodegeneration or rather are symptoms of other disease mechanisms (Castellani, Lee, Zhu, Perry, &

[1] Original quote in British English.

Smith, 2008). In contrast, VD emerges from global or focal effects of cerebrovascular lesions evoked by cerebrovascular diseases such as atherosclerosis, small vessel disease, or cerebral amyloid angiopathy. More detailed, infarction, hemorrhage, white matter lesions, or microvascular ischemia lead to brain tissue destructions of different sizes and locations (Haaland, 2015; Walker, McAleese, Erskine, & Attems, 2019). Accordingly, there are several subtypes of VD including cortical VD/multi-infarct dementia, strategic infarct dementia, and subcortical VD/small vessel dementia (Rockwood, 2002). Mixed dementia occurs when the disease pathology includes cerebrovascular and neuropathological causes (e.g. combination of VD and AD; Walker et al., 2019). In the following, this thesis distinguishes between AD and non-AD etiologies (e.g. VD, mixed dementia, and other degenerative etiologies), when reporting research findings.

Independent of etiology, dementia is characterized by brain atrophy and destructions of brain tissue, which are related to various cognitive, motor, and functional impairments (Albert, 2011; Bennett et al., 2012; Parihar, Mahoney, & Verghese, 2013; Valkanova & Ebmeier, 2017). These vary across different severities and etiologies of dementia with respect to time and extent of occurrence (Cohen-Mansfield, 2000). The following sections summarize characteristic cognitive, motor, and functional impairments associated with dementia considering different severities and etiologies.

2.1.1 Dementia and cognitive impairments

Cognitive impairments are the hallmark of dementia (Albert, 2011; Waldemar et al., 2007). Dependent on the etiology and severity of dementia, several cognitive functions are affected (Weintraub, Wicklund, & Salmon, 2012). Memory deficits are the most prominent symptoms of AD (Stopford, Thompson, Neary, Richardson, & Snowden, 2012; Weintraub et al., 2012). In early disease stages, AD is characterized by a progressive decline of episodic memory. With advancing severity other dimensions of memory, e.g. semantic memory, are affected and additional cognitive decline in language, executive function, attention, working memory, and visuospatial function occurs (Lindeboom & Weinstein, 2004; Perry & Hodges, 1999; Weintraub et al., 2012). In VD, cognitive impairments vary depending on localization and size of cerebrovascular lesions. Despite no single characteristic neurological profile exists for VD, a decline of executive function is common (Desmond, 2004a; A. Y. Lee, 2011). Moreover, VD is often attributed to stepwise deterioration (Rockwood, 2002). A closer look at VD sub-

types, however, shows different patterns of cognitive impairments. Cortical VD is characterized by abrupt onset of cognitive impairments, language deficits, and loss of executive function. In mild severities, it is very weak and typically progresses stepwise. Compared to that, cognitive impairments considerably differ among strategic infarct dementia, but often include a decline of memory, as various cortical or subcortical areas are affected. In contrast, impaired executive function is the most prominent symptom of subcortical VD, which is commonly located in specific areas in the prefrontal subcortical circuit. Additionally, memory deficits less pronounced than in AD occur (Desmond, 2004b; Rockwood, 2002). In mixed dementia, AD and VD often coexist (Rockwood, 2002). Related to the combination of different pathophysiologies, additive or synergistic effects of both etiologies on cognitive impairments have been suggested (Attems & Jellinger, 2014).

Summarizing these cognitive impairments, dementia is typically characterized by progressive decline affecting memory, language, executive function, attention, working memory, and visuospatial function (Beauchet et al., 2008; Weintraub et al., 2012). Estimating impacts of these impairments when performing studies with IWD, it is worth to take a more detailed look at involved cognitive domains. Table 1 includes short descriptions of cognitive domains impaired in IWD and related impacts.

Table 1. Description of cognitive domains impaired in individuals with dementia and related impacts

	Description of cognitive domains	**Impacts of cognitive impairments**
Memory	*Episodic memory*: "recollection of past events in their context of time and space" (Lindeboom & Weinstein, 2004, p. 84)	Impaired ability to learn and remember new information (Weintraub et al., 2012)
	Semantic memory: knowledge of concepts and facts (Lindeboom & Weinstein, 2004)	Negative effects on language, as semantic memory also includes meanings of words (Weintraub et al., 2012)
Language	Ability to express thoughts into words or symbols for communication purposes including language production (e.g. speaking, writing) and comprehension (e.g. comprehending, reading; Tang-Wai & Graham, 2008)	Impacts on speech expression, naming, and comprehension (Lindeboom & Weinstein, 2004) affecting word-finding, verbal fluency, object naming, semantic categorization, sentence comprehension, or discourse cohesion (Kempler & Goral, 2008; Tang-Wai & Graham, 2008; Weintraub et al., 2012)
Executive function	Higher cognitive capacities including goal formulation, initiation, planning, organizing, and regulation of goal-directed behavior (Lezak, 2012; Lindeboom & Weinstein, 2004; Rockwood, 2002)	Disturbed concept formation, affected mental manipulation of information, impaired problem solving, inadequate cue-directed behavior, apathy, loss of initiative or conversely disinhibition and impulsivity, and perseveration (Lindeboom & Weinstein, 2004; Perry & Hodges, 1999; Weintraub et al., 2012)

Attention	Closely related to executive function (Lindeboom & Weinstein, 2004; Weintraub et al., 2012) *Selective attention*: focusing on a single relevant stimulus, while ignoring irrelevant ones (Perry & Hodges, 1999) *Divided attention*: sharing attention between two simultaneous stimuli (Perry & Hodges, 1999) *Sustained attention*: maintaining focus over an extended period (Perry & Hodges, 1999)	Inability to concentrate, easy distractibility, and difficulties in complex tasks requiring effective allocation of attentional resources (e.g. dual tasks), disengagement, or shifting of attention (Perry & Hodges, 1999; Weintraub et al., 2012)
Working memory	Closely related to executive function (Lindeboom & Weinstein, 2004; Weintraub et al., 2012) Processing system temporarily storing information of immediate focus, which supports human thinking by connecting perception, long-term memory, and action (Baddeley, 2003)	Problems with memory span, digit reversal, spelling, calculation, and encoding of lists of stimuli in selective or divided attention tasks (Stopford et al., 2012)
Visuo-spatial function	Multi-faceted set of functions including orienting attention, appreciating positions of stimulus-objects in space, integrating objects into a coherent spatial framework, mental operations involving spatial concepts and navigation learning (Geldmacher, 2003; Possin, 2010)	Problems of spatial thinking, impaired construction abilities, deficits in spatial orientation, especially in less familiar surroundings, and forgetting where personal items have been placed (Karantzoulis & Galvin, 2011; Kirova, Bays, & Lagalwar, 2015; Lindeboom & Weinstein, 2004; Weintraub et al., 2012)

Impacts of cognitive impairments in dementia reflect challenges related to everyday life and in dealing with IWD. Based on these impacts, especially impairments in memory, language, executive function, and attention may affect investigations with IWD. It is important to appropriately consider these impairments, particularly with respect to adequate assessments and interventions.

2.1.2 Dementia and motor impairments

Compared to cognitive impairments, motor impairments are less frequently considered associated with dementia. The ICD-10 criteria for dementia do not include indications for motor impairments. Other diagnostic criteria mention motor impairments only for some etiologies or with advanced severity. In AD, motor impairments, for instance, are frequently regarded as late symptoms not affecting mild severities (McKhann et al., 1984; Pettersson et al., 2005). However, there are several studies showing impaired motor performance in various etiologies and even in mild dementia (Pettersson et al., 2005).

Generally, these studies show that IWD are characterized by poorer motor performance compared to cognitively unimpaired older adults and individuals with mild cognitive impairment (MCI) (Allan, Ballard, Burn, & Kenny, 2005; Eggermont et al., 2010; Gras et al., 2015; Kato-Narita, Nitrini, & Radanovic, 2011; Leandri et al., 2009; Pettersson et al., 2005; Pettersson, Engardt, & Wahlund, 2002; Suttanon et al., 2012; Tangen, Engedal, Bergland, Moger, & Mengshoel, 2014). With respect to the severity of dementia, decreasing motor performance is observed with increasing cognitive impairments (Coelho et al., 2012; Leandri et al., 2009; McGough et al., 2013; Tangen et al., 2014). Moreover, differences in motor performance between different etiologies are reported. In detail, individuals with non-AD, perform worse than those with AD (Allan et al., 2005; Pettersson et al., 2005). With respect to different domains of motor performance, Table 2 summarizes available findings for balance, mobility, strength, endurance, and functional performance comprising several motor domains by exemplarily showing the results of previous cross-sectional studies.

Table 2. Summary of motor impairments in individuals with dementia observed in previous studies considering the severity and etiology of dementia

	Observed motor impairments	**Influence of severity and etiology of dementia**
Balance	Poorer balance performance is consistently reported for IWD compared to cognitively unimpaired older adults and individuals with MCI (Allan et al., 2005; Gras et al., 2015; Kato-Narita et al., 2011; Leandri et al., 2009; Pettersson et al., 2002; Pettersson et al., 2005; Suttanon et al., 2012; Tangen et al., 2014).	Greater balance impairments are associated with lower cognitive performance (Kato-Narita et al., 2011; Leandri et al., 2009; Mazoteras Muñoz et al., 2010; McGough et al., 2013; Tangen et al., 2014). Balance impairments also occur in mild dementia (Allan et al., 2005; Gras et al., 2015; Leandri et al., 2009; Pettersson et al., 2002; Tangen et al., 2014). Individuals with non-AD show more balance impairments than individuals with AD (Allan et al., 2005; Pettersson et al., 2005).
Mobility	Studies investigating mobility impairments in IWD frequently focus on spatio-temporal gait parameters. Appropriate findings are summarized in Chapter 2.1.3 and not considered herein. IWD show worse mobility performance than cognitively unimpaired older adults and individuals with MCI (Eggermont et al., 2010; Gras et al., 2015; Pettersson et al., 2005).	Worse mobility performance and greater mobility impairments, respectively are observed in moderate compared to mild AD (Coelho et al., 2012) and non-AD compared to AD (Pettersson et al., 2005).

Strength	It is concluded that muscle weakness is common in IWD (Nakayama et al., 2019). No differences in functional lower-limb strength are observed between IWD and cognitively unimpaired older adults and individuals with MCI. However, a high amount of missing values may have resulted in insufficient power (Eggermont et al., 2010). Different motor strategies for sit-to-stand movements are reported for IWD compared to cognitively unimpaired older adults (Manckoundia, Mourey, Pfitzenmeyer, & Papaxanthis, 2006).	No investigations identified.
Endurance	63.6 % of IWD have poor cardiorespiratory endurance, which however is not further specified (Arshinta, Fitriana, Adikusuma, Rohaedi, & Putri, 2018). No investigations comparing the endurance performance of IWD and cognitively unimpaired older adults/individuals with MCI identified.	No investigations identified.
Functional performance comprising several motor domains	No investigations identified.	A statistically significant correlation between functional performance and global cognition reflects greater functional performance impairments with increasing cognitive impairments in IWD (McGough et al., 2013).

IWD: individuals with dementia, MCI: mild cognitive impairment

Summarizing the findings of previous studies shows that IWD are frequently affected by motor impairments. These impairments and related declines in motor performance exceed those observed in healthy aging. While balance and mobility impairments in IWD are reported in several studies, information on strength, endurance, and functional performance is rare.

2.1.3 Gait characteristics associated with dementia as an example for functional impairments

Gait characteristics commonly change with increasing age (Jahn, Zwergal, & Schniepp, 2010). For instance, walking speed decreases by 1 % per year from the age of 60 (Ashton-Miller, 2005). Gait impairments exist when there is a “demonstrable gait abnormality beyond the normal age-related slowing” (Jahn et al., 2010, p. 307). Gait impairments are common in IWD (Annweiler et al., 2012; van Iersel, Hoefsloot, Munneke, Bloem, & Olde Rikkert, 2004) and affect more than 50 % (Allali & Verghese, 2017).

Gait impairments in dementia is a frequently considered research topic. Appropriate reviews and research articles, combine main findings and thus give a good overview of the current knowledge. Typical gait impairments associated with dementia comprise decreased walking speed, shortened stride/step length, increased double support time, and enhanced step to step variability (Alexander & Hausdorff, 2008; Beauchet et al., 2008; Coelho et al., 2012; Dorfman, Mirelman, Hausdorff, & Giladi, 2014; Ijmker & Lamoth, 2012; Scherder et al., 2007; Sheridan, Solomont, Kowall, & Hausdorff, 2003; Valkanova & Ebmeier, 2017; van Iersel et al., 2004). Similar to motor impairments, gait impairments in IWD are more pronounced than in cognitively unimpaired older adults and in individuals with MCI (Amboni et al., 2013; Beauchet et al., 2008; Dorfman et al., 2014; Ijmker & Lamoth, 2012; Mc Ardle et al., 2017; Scherder et al., 2007; Sheridan et al., 2003; Valkanova & Ebmeier, 2017; van Iersel et al., 2004), and increase with advancing severity of dementia (Allali & Verghese, 2017; Amboni et al., 2013; Annweiler et al., 2012; Beauchet et al., 2008; Scherder et al., 2007; Valkanova & Ebmeier, 2017; van Iersel et al., 2004). Moreover, gait impairments are greater in non-AD, especially in VD, compared to AD (Allali & Verghese, 2017; Beauchet et al., 2018; Dorfman et al., 2014; Mc Ardle et al., 2017; Valkanova & Ebmeier, 2017; van Iersel et al., 2004). Nevertheless, Scherder et al. (2007) conclude based on the main findings of several studies that gait impairments are present in all etiologies of dementia and even affect mild severities.

Besides assessing gait in single task conditions, several studies also apply dual tasks. Herein, they examine gait performance while participants perform an additional cognitive task, such as counting backwards or naming animals (Kressig & Beauchet, 2006). Findings on gait impairments of studies using dual tasks are comparable to those observed in single task conditions. However, dual tasks enhance the sensitivity of gait assessments. Some studies, are only able to detect gait impairments during dual task conditions (Valkanova & Ebmeier, 2017). In line with this, gait impairments are more pronounced than in single task conditions (Amboni et al., 2013; Valkanova & Ebmeier, 2017). Moreover, dual task costs are larger in IWD than in cognitively unimpaired older adults and related to the severity of dementia (Amboni et al., 2013).

Aiming to get a closer insight into gait impairments of IWD, findings of individual studies are exemplarily summarized in Table 3. Several studies investigate numerous spatio-temporal gait parameters and compare findings in different severities and etiologies of dementia.

Table 3. Summary of gait impairments in individuals with dementia observed in previous studies considering the severity and etiology of dementia

	Observed values of spatiotemporal gait parameters (considered studies)	**Observed gait impairments**	**Influence of severity and etiology of dementia**	**Findings on gait impairments in dual task conditions**
Walking speed	*IWD*: 36.0-110.9 cm/s *Cognitively unimpaired older adults*: 70.0-149.0 cm/s *MCI*: 75.0-122.0 cm/s (Allali et al., 2016; Beauchet et al., 2018; Beauchet, Allali, Launay, Herrmann, & Annweiler, 2013; Cadore et al., 2015; Eggermont et al., 2010; Gillain et al., 2009; Gras et al., 2015; Ijmker & Lamoth, 2012; Lamoth et al., 2011; Maquet et al., 2010; Merory, Wittwer, Rowe, & Webster, 2007; Muir et al., 2012; Nadkarni, Mawji, McIlroy, & Black, 2009; Nakamura et al., 1997; Rucco et al., 2017; Theill, Martin, Schumacher, Bridenbaugh, & Kressig, 2011; Visser, 1983; Webster, Merory, & Wittwer, 2006; Wittwer, Webster, & Menz, 2010)	The majority of previous studies report statistically significant lower walking speed in IWD compared to cognitively unimpaired older adults and individuals with MCI (except Lamoth et al., 2011; Muir et al., 2012).	Statistically significant lower walking speed is observed in moderate compared to mild dementia (Beauchet et al., 2018) or only in moderate to severe dementia compared to cognitively unimpaired older adults, but not in mild severities (Nakamura et al., 1997) as well as in non-AD compared to AD (Allali et al., 2016) or only for non-AD compared to cognitively unimpaired older adults (Rucco et al., 2017).	With one exception (Lamoth et al., 2011), all studies observe statistically significant slower walking speed in IWD compared to cognitively unimpaired older adults and individuals with MCI during dual task conditions (Gillain et al., 2009; Ijmker & Lamoth, 2012; Lamoth et al., 2011; Muir et al., 2012; Rucco et al., 2017; Theill et al., 2011).
Cadence and stride/step frequency	Different reported values (steps/min, strides/s, Hz) are not directly comparable and thus are not presented. (Gillain et al., 2009; Lamoth et al., 2011; Merory et al., 2007; Nadkarni, Mawji et al., 2009; Rucco et al., 2017; Visser, 1983; Wittwer et al., 2010)	Most studies observe no statistically significant differences for cadence and stride/step frequency (Gillain et al., 2009; Lamoth et al., 2011; Merory et al., 2007; Wittwer et al., 2010).	Statistically significant lower cadence is reported for individuals with non-AD compared to cognitively unimpaired older adults, but not for individuals with AD (Rucco et al., 2017).	Cadence assessed in dual task condition is statistically significantly lower in IWD compared to cognitively unimpaired older adults (Rucco et al., 2017).

Stride/step length	*IWD*: 49.0-119.6 cm/43.0-62.0 cm *Cognitively unimpaired older adults*: 103.0-141.0 cm/58.4-77.0 cm *MCI*: 111.3-136.0 cm/N/A (Allali et al., 2016; Beauchet et al., 2018; Gillain et al., 2009; Merory et al., 2007; Nadkarni, Mawji et al., 2009; Nakamura et al., 1997; Webster et al., 2006; Wittwer et al., 2010)	With one exception (Wittwer et al., 2010), all studies observe statistically significant shorter stride/step length in IWD compared to cognitively unimpaired older adults and individuals with MCI.	Statistically significant lower stride length is reported for moderate compared to mild severity of dementia (Beauchet et al., 2018; Nakamura et al., 1997) and non-AD compared to cognitively unimpaired older adults or individuals with AD (Allali et al., 2016; Merory et al., 2007; Rucco et al., 2017).	In dual task conditions, statistically significant shorter stride length is observed for IWD compared to cognitively unimpaired older adults and individuals with MCI (Merory et al., 2007; Rucco et al., 2017).
Stride/step time	*IWD*: 1.1-1.3 s/0.58 s *Cognitively unimpaired older adults*: 1.0-1.2 s/0.55 s *MCI*: 1.1-1.2 s/N/A (Allali et al., 2016; Beauchet et al., 2018; Choi et al., 2011; Ijmker & Lamoth, 2012; Lamoth et al., 2011; Muir et al., 2012; Nadkarni, Mawji et al., 2009; Rucco et al., 2017; Wittwer et al., 2010)	Results are inconsistent, with one third of considered studies observing statistically significant larger stride/step time in IWD compared to cognitively unimpaired older adults and individuals with MCI (Beauchet et al., 2018; Ijmker & Lamoth, 2012; Nadkarni, Mawji et al., 2009).	Stride time is statistically significantly larger in moderate compared to mild severity of dementia (Beauchet et al., 2018) and non-AD compared to cognitively unimpaired older adults, but not in AD (Rucco et al., 2017).	All but one (Lamoth et al., 2011) studies report statistically significant larger stride time in IWD compared to cognitively unimpaired older adults during dual task conditions (Ijmker & Lamoth, 2012; Lamoth et al., 2011; Muir et al., 2012; Rucco et al., 2017).
Stride/step width	*IWD*: 9.2-12.3 cm *Cognitively unimpaired older adults*: 8.9-10.0 cm *MCI*: 9.8-10.0 cm (Allali et al., 2016; Beauchet et al., 2018; Merory et al., 2007; Nadkarni, Mawji et al., 2009; Rucco et al., 2017; Webster et al., 2006; Wittwer et al., 2010)	The majority of studies does not observe statistically significant differences in stride/step width between IWD and cognitively unimpaired older adults or individuals with MCI (Merory et al., 2007; Nadkarni, Mawji et al., 2009; Webster et al., 2006; Wittwer et al., 2010).	Stride width is statistically significantly larger in moderate than in mild severity of dementia (Beauchet et al., 2018) and in non-AD compared to AD (Allali et al., 2016). In contrast, Rucco et al. (2017) observe a statistically significant larger stride width compared to cognitively unimpaired older adults only in AD, but not in non-AD.	No investigations identified.

Stance and swing time	*IWD*: 0.71-0.91 s and 0.40-0.41 s *Cognitively unimpaired older adults*: 0.52-0.76 s and 0.41 s *MCI*: 0.79-0.81 s and 0.42 s (Allali et al., 2016; Beauchet et al., 2018; Gras et al., 2015; Rucco et al., 2017)	Half of the considered studies report statistically significant higher stance time and/or lower swing time in IWD compared to cognitively unimpaired older adults and individuals with MCI (Beauchet et al., 2018; Gras et al., 2015).	Besides statistically significant higher stance time compared to cognitively unimpaired older adults in individuals with non-AD, but not in individuals with AD (Rucco et al., 2017), no differences with respect to severity and etiology of dementia are observed for stance and swing time (Allali et al., 2016; Beauchet et al., 2018; Rucco et al., 2017).	No investigations identified.
Double and single support (time/percent of cycle)	*IWD*: 0.35-0.51 s/18.8-24.2 % and 0.40-0.41 s/N/A *Cognitively unimpaired older adults*: 0.21-0.35 s/11.0-24.7 % and 0.41 s/N/A *MCI*: 0.37-0.38 s/N/A and 0.42 s/N/A (Allali et al., 2016; Beauchet et al., 2018; Merory et al., 2007; Nadkarni, Mawji et al., 2009; Nakamura et al., 1997; Rucco et al., 2017; Visser, 1983; Wittwer et al., 2010)	With single exceptions (Wittwer et al., 2010) statistically significant higher double support and lower single support values are reported for IWD compared to cognitively unimpaired older adults and individuals with MCI.	Statistically significant smaller double support times are shown for mild compared to moderate dementia and in moderate to severe dementia compared to cognitively unimpaired older adults, while findings comparing mild dementia and cognitively unimpaired older adults are inconsistent and no differences are reported for single support time (Beauchet et al., 2018; Nakamura et al., 1997). No differences with respect to etiologies are observed (Allali et al., 2016; Merory et al., 2007; Rucco et al., 2017).	No investigations identified.

IWD: individuals with dementia, MCI: mild cognitive impairment, N/A: not available

Besides spatiotemporal gait parameters, step-to-step variability is frequently examined in previous studies (Allali et al., 2016; Beauchet et al., 2013; Beauchet et al., 2018; Choi et al., 2011; Ijmker & Lamoth, 2012; Lamoth et al., 2011; Muir et al., 2012; Nakamura et al., 1997; Rucco et al., 2017; Visser, 1983; Webster et al., 2006; Wittwer et al., 2010). Large heterogeneity of assessed parameters exists and variability differs dependent on assessed parameters. Inconsistent findings between studies compromise drawing conclusions. For walking/stride speed, stride length, and stride/step time previous studies predominately report statistically significant higher variability in IWD compared to cognitively unimpaired older adults (Beauchet et al., 2013; Beauchet et al., 2018; Choi et al., 2011; Ijmker & Lamoth, 2012; Nakamura et al., 1997; Webster et al., 2006; Wittwer et al., 2010).

Based on previous studies gait impairments including reduced walking speed, shortened stride/step length, enhanced stance time, and increased double support can be assumed in IWD. Findings on other gait parameters are inconsistent. Moreover, available studies do not allow profound conclusions on the influence of severity and etiology of dementia, as most focus on AD and do not distinguish between different severities of dementia. Nevertheless, there are some indications for increased gait impairments in advanced severities of dementia and non-AD. These findings are in line with those reported in previous reviews (Mc Ardle et al., 2017; van Iersel et al., 2004).

2.2 Understanding gait impairments in individuals with dementia

Often, gait is considered primarily as an “automated, over-learned, rhythmic motor task” (Hausdorff, Yogev, Springer, Simon, & Giladi, 2005, p. 541). However, several investigations show the importance of cognitive function in gait (Dorfman et al., 2014; R. Morris, Lord, Bunce, Burn, & Rochester, 2016). More detailed, gait involves axial musculature, balance, and movements of bilateral upper and lower extremities, but also cognitive control of these motor processes, as well as the incorporation of sensory feedback (J. A. Cohen, Verghese, & Zwerling, 2016). Accordingly, gait is a complex non-linear process, which requires input from the cerebellum, the motor cortex, the basal ganglia, as well as visual, vestibular, and proprioceptive feedback (Hausdorff, 2007).

This chapter considers gait and gait impairments in the context of different causes, relations, and influences. At first, gait impairments and underlying structural changes

in the brain are briefly discussed. Afterward, knowledge about relations between cognitive and gait performance in IWD are summarized. Finally, motor performance and its influence on gait are regarded.

2.2.1 Gait performance and brain changes in individuals with dementia

There are various pathologic processes in the brain associated with dementia and gait impairments such as ischemia, inflammation, neurodegeneration, and cerebral white matter hyperintensities (Parihar et al., 2013; Rosano et al., 2005; Valkanova & Ebmeier, 2017). Cortical regions like temporal and frontal lobes comprising primary motor areas and the hippocampus, frontal subcortical circuits encompassing basal ganglia, and the cerebellum are affected by such pathologic processes. These brain structures are involved in movement planning, initiation, processing of information, and gait control (Annweiler et al., 2012; Beauchet et al., 2008; Beauchet, Launay, Sekhon, Montembeault, & Allali, 2019; Parihar et al., 2013). There seems to be a linked alteration between structural changes and gait impairments (Alexander & Hausdorff, 2008; Mc Ardle et al., 2017). Recent reviews give a comprehensive overview of relations between pathologic processes and gait impairments in cognitively unimpaired older adults and IWD (see Annweiler et al., 2012; Tian et al., 2017; Valkanova & Ebmeier, 2017). Nevertheless, studies with IWD are rare and most of them focus on individuals with AD. Thus, only isolated findings of previous investigations can be presented. These include relations between increased gait impairments and lower cerebral blood flow in prefrontal cortex and basal ganglia (Nakamura et al., 1997), statistically nonsignificant trends for a negative association between hippocampal volume and stride time variability (Beauchet et al., 2019), and statistically significant correlations between overall white matter hyperintensities and stride length/walking speed (Nadkarni, McIlroy, Mawji, & Black, 2009).

Despite these rare findings, causal relations between structural brain changes and gait impairments in IWD can be assumed based on several theoretical indications presented above. Additionally, these brain changes do not only induce gait impairments but also cognitive impairments, which might explain associations between gait and cognitive performance in IWD (Alexander & Hausdorff, 2008; Ferrer, 2010; Mc Ardle et al., 2017; Nelson et al., 2012).

2.2.2 Gait and cognitive performance in individuals with dementia

A close connection between gait and cognitive performance is postulated (Scherder et al., 2007). Such associations exist in cognitively unimpaired older adults and IWD (Alexander & Hausdorff, 2008). Moreover, it is suggested that gait impairments in IWD may be a specific sign of disease-related cognitive decline (Annweiler et al., 2012; Beauchet et al., 2008). Based on the consequences and high prevalence of gait impairments and cognitive decline in older adults and IWD, research focusing on relations between gait and cognition has gained increasing attention in the last decade (Amboni et al., 2013).

Focusing on IWD, relations between gait performance and several cognitive functions are observed, especially for executive function, attention, and working memory (Allali et al., 2016; Amboni et al., 2013; Montero-Odasso, Verghese, Beauchet, & Hausdorff, 2012; R. Morris et al., 2016; Valkanova & Ebmeier, 2017). In line with this, there is evidence that safe and effective gait requires high-level cognitive input involving executive function, attention, and memory (Hausdorff et al., 2005; Scherder et al., 2007; Valkanova & Ebmeier, 2017; Yogev-Seligmann, Hausdorff, & Giladi, 2008). Table 4 exemplarily summarizes, substantial findings of previous studies investigating associations between gait and cognitive performance in IWD.

Table 4. Summary of relations between gait and cognitive performance in individuals with dementia observed in previous studies

	Relation between gait and cognitive performance in single task conditions	**Relation between gait and cognitive performance in dual task conditions**
Global cognition	Moderate to high statistically significant correlations of global cognition with walking speed, cadence, step length, stride time, and stride time variability are observed in IWD (Bruce-Keller, Brouillette, Tudor-Locke, Foil, Gahan, Nye et al., 2012; Ijmker & Lamoth, 2012; Lamoth et al., 2011).	Moderate to high statistically significant correlations of global cognition with walking speed and changes in stride-to-stride variability between single and dual task, but not with stride length and cadence are reported in IWD for dual task conditions (Bruce-Keller, Brouillette, Tudor-Locke, Foil, Gahan, Correa et al., 2012; Sheridan et al., 2003).
Executive function	Moderate statistically significant correlations of executive function with cadence and step length, inconsistent findings for walking speed, and no correlations with stride time and stride time variability are shown in IWD (Bruce-Keller, Brouillette, Tudor-Locke, Foil, Gahan, Nye et al., 2012; Lamoth et al., 2011).	Moderate to high statistically significant correlations of executive function with walking speed, stride speed, cadence, and changes in stride-to-stride variability between single and dual task, but not with stride length are observed in IWD for dual task conditions (Bruce-Keller, Brouillette, Tudor-Locke, Foil, Gahan, Correa et al., 2012; Coelho et al., 2012).

Attention	No correlations of attention with gait performance in a sample of IWD and cognitively unimpaired older adults are reported (Ijmker & Lamoth, 2012).	No investigations identified.
Working memory	No correlations with gait performance are shown in a sample of IWD and cognitively unimpaired older adults (Ijmker & Lamoth, 2012).	No investigations identified.
Memory	No correlations of memory with walking speed, stride time, and stride time variability are observed in IWD (Lamoth et al., 2011).	No investigations identified.
Verbal fluency	Moderate to high statistically significant correlations of verbal fluency with walking speed, cadence, step length, stride time, and stride time variability are reported in IWD (Bruce-Keller, Brouillette, Tudor-Locke, Foil, Gahan, Nye et al., 2012; Ijmker & Lamoth, 2012; Lamoth et al., 2011).	Moderate statistically significant correlations of verbal fluency with changes in stride-to-stride variability between single and dual task, but no correlation with spatiotemporal gait parameters are shown in IWD for dual task conditions (Bruce-Keller, Brouillette, Tudor-Locke, Foil, Gahan, Correa et al., 2012; Sheridan et al., 2003).

IWD: individuals with dementia

Summarizing these findings, previous studies show statistically significant associations between lower gait performance and decreased cognitive performance in global cognition, executive function, and verbal fluency. Moreover, reviews and overview article strongly indicate that gait performance in IWD is related to attention and working memory. In contrast, the examination of exemplary studies does not support this associations. However, only one study focusing on these two cognitive domains was identified, which includes IWD and cognitively unimpaired older adults. This may not reflect actual associations. Moreover, no correlations with memory are observed, but again only one study was considered.

2.2.3 Gait and motor performance in individuals with dementia

Besides cognition, sensory and motor systems are important for safe and effective gait (Callisaya et al., 2009). Associations between impaired gait performance, e.g. slower walking speed or reduced step length, and motor performance, e.g. strength, balance, or reaction time, have been suggested in cognitively unimpaired older adults and IWD (Alexander, 1996; Bruce-Keller, Brouillette, Tudor-Locke, Foil, Gahan, Nye et al., 2012; Callisaya et al., 2009; J. A. Cohen et al., 2016; Suzuki et al., 2012; Tiedemann, Sherrington, & Lord, 2005). In both populations, a decrease in motor performance is common and may contribute to gait impairments (Callisaya et al., 2009; J. A. Cohen et al., 2016; Tiedemann et al., 2005). This chapter exemplarily presents studies examining

the relationship between gait and motor performance (see Table 5). As research in IWD primarily focuses on cognitive influences, it starts with findings of older adults to build a fundamental basis, which is supplement by specific results in IWD.

Table 5. Summary of relations between gait and motor performance in cognitively unimpaired older adults and individuals with dementia observed in previous studies

	Motor domain	Relations between gait and motor performance
Cognitively unimpaired older adults	**Strength**	Better quadriceps strength is a predictor of faster walking speed and reduced double support phase (only in women), and moreover is associated with higher step length, faster cadence, and gait variability measures (Callisaya et al., 2009; Callisaya, Blizzard, McGinley, Schmidt, & Srikanth, 2010).
	Balance	Better postural sway is a predictor of faster walking speed (only in men) and reduced double support phase, and moreover is associated with higher step length, smaller step width, and gait variability measures (Callisaya et al., 2009; Callisaya et al., 2010).
	Reaction time	Better reaction time is a predictor of faster walking speed and reduced double support phase, and moreover is associated with higher step length, faster cadence, and gait variability measures (Callisaya et al., 2009; Callisaya et al., 2010).
Individuals with dementia	**Strength**	Knee extension strength is a statistically significant predictor of gait performance (Suzuki et al., 2009; Suzuki et al., 2012). Moreover, statistically significant correlations between knee extension strength and independence level of gait are observed (Nakayama et al., 2019).
	Balance	Positive associations between walking speed/cadence and balance are reported, while associations between double support phase/stride length variability/stride time variability and balance are negative (McGough et al., 2013).
	Functional performance	Positive associations between walking speed/cadence and functional performance (including balance, mobility, and strength) are reported, while associations between double support phase/stride length variability/stride time variability and functional performance are negative (McGough et al., 2013).

Studies investigating cognitively unimpaired older adults and IWD show similar findings. With respect to motor performance, especially lower limb strength and balance performance seem to be related to gait performance. Consistently, associations differ depending on the gait parameter considered. In line with this, Alexander (1996) suggests that changes related to musculoskeletal function, e.g. strength, potentially limit gait parameters determined by energy expenditure, such as walking speed or stride/step length, while balance deteriorations rather may affect postural gait parameters like double support time.

Summary

Theoretical foundation for establishing a high-quality methodological approach

- Cognitive impairments especially in memory, language, executive function, and attention, as well as motor impairments in balance, mobility, and strength, frequently occur in IWD and need to be considered when selecting motor assessments and developing physical interventions specifically tailored to IWD.
- Established using the example gait, there is a close relationship between motor and cognitive performance in IWD. In general, safe and effective gait requires higher cognitive input and is dependent on several motor functions. Moreover, associations between gait and cognitive performance support the application of combined physical activity and cognitive interventions.

Theoretical foundation for performing a high-quality RCT investigating the effectiveness of physical activity on motor performance in IWD

- In IWD, impairments in balance, mobility, and strength occur, which can potentially be influenced by physical activity.

Theoretical foundation for performing a high-quality RCT investigating the effectiveness of physical activity on gait performance, characteristics of responders, and impacts of changes in underlying motor and cognitive performance on changes in gait performance

- Reduced walking speed, shortened stride/step length, enhanced stance time, and increased double support are characteristic gait impairments in IWD. Inconsistent findings are available for stride/step time. These spatiotemporal gait parameters may be sensitive to physical activity.
- There are several differences in cognitive, motor, and gait performance between individuals with varying severities and etiologies of dementia. With advancing cognitive impairments, cognitive, motor, and gait performance decreases in IWD. Furthermore, cognitive impairments differ related to underlying pathophysiology. Additionally, greater motor and gait impairments are reported for individuals with non-AD compared to individuals with AD. These differences associated with the severity and etiology of dementia potentially influence the effectiveness of physical activity.
- Several associations of spatiotemporal gait parameters with cognitive and motor performance are reported in IWD. Based on theoretical considerations and empirical findings these include executive function, attention, and working memory, as well as strength, balance, and functional performance. Thus, changes in these cognitive and motor functions may contribute to explain intervention-induced changes in gait performance.

3 Current state of research

Physical activity is discussed as an important therapeutic strategy for dementia (Ahlskog, Geda, Graff-Radford, & Petersen, 2011). There are various studies and reviews investigating the effectiveness of physical activity on motor and cognitive performance in IWD (e.g. Blankevoort et al., 2010; Bossers et al., 2015; Brett et al., 2016; Groot et al., 2016; Schwenk, Dutzi et al., 2014). This chapter aims to summarize the current state of research focusing on motor performance in general and specifically on gait as functional performance related to motor and cognitive performance, frequently impaired in IWD. It is based on two systematic searches supplemented by findings of recent reviews (2010 or later). The first search was performed to identify RCT investigating the effectiveness of physical activity on motor performance in IWD[2]. The second search focuses on studies examining the effectiveness of physical activity on spatiotemporal gait parameters[3]. Since the number of RCT is limited in this field of research the second search includes studies of all designs applying inferential statistics. Subsequently, a short summary of the effectiveness of physical activity on cognitive performance in IWD is given. This summary only considers the findings of recent reviews. Finally, Chapter 3 ceases with an overview discussing the limitations of previous studies, which are important to consider when concluding on evidence.

3.1 Effectiveness of physical activity on motor performance in individuals with dementia

The current state of research on the effectiveness of physical activity on motor performance is based on 46 RCT and eight recent reviews. RCT assess the effects of different physical activity interventions on balance, mobility, strength, endurance, flexibility, and functional performance, while recent reviews summarize findings of previous investigations.

[2] The first systematic search was performed for manuscript II (Trautwein, Maurus, Barisch-Fritz, Hadzic, and Woll (2019)). Further details are given in chapter 5.2. This systematic review does not consider the effectiveness of physical activity on motor performance, but focuses on motor assessments applied in RCT investigating effectiveness of physical activity. Accordingly, RCT were identified to summarize applied motor assessments.

[3] The second systematic search belongs to a systematic review assessing the effectiveness of physical activity on spatiotemporal gait parameters in IWD (Trautwein, Hadzic, Barisch-Fritz, Maurus, and Woll (in prep.)). Further details are provided at the International Prospective Register of Systematic Reviews (registration number: CRD42018106370).

Seven recent reviews conclude that physical activity generally has positive effects on motor performance in IWD (Blankevoort et al., 2010; Brett et al., 2016; Hernández et al., 2015; Lam, Huang et al., 2018; Pitkälä, Savikko, Pöysti, Strandberg, & Laakkonen, 2013; Potter, Ellard, Rees, & Thorogood, 2011; Suttanon et al., 2010). However, statistically significant positive effects are not found for all motor domains and are dependent on applied interventions (Blankevoort et al., 2010; Brett et al., 2016; Lam, Huang et al., 2018; Pitkälä, Savikko et al., 2013; Potter et al., 2011; Suttanon et al., 2010). Accordingly, derived evidence is not rated consistently throughout recent reviews, ranging from limited (Suttanon et al., 2010) to some (Potter et al., 2011) and low to moderate grade of evidence (Pitkälä, Savikko et al., 2013), respectively.

As the effectiveness of physical activity in IWD mostly seems to be depending on the motor domain and applied interventions, these two aspects are taken into account more detailed. Therefore, findings of identified RCT and recent reviews are summarized.

3.1.1 Balance

Balance is examined in 27 RCT using different balance assessments. For static balance, seven of ten RCT observe statistically significant improvements with small to large effect sizes (Arcoverde et al., 2014; Bossers et al., 2015; Burgener, Yang, Gilbert, & Marsh-Yant, 2008; Kampragkou, Iakovidis, Kampragkou, & Kellis, 2017; Miu, Szeto, & Mak, 2008; Suttanon et al., 2013; Vreugdenhil, Cannell, Davies, & Razay, 2012 vs. Netz, Axelrad, & Argov, 2007; Rolland et al., 2007; Wesson et al., 2013). Moreover, statistically significant improvements are found in four of five RCT assessing postural sway (Schwenk, Dutzi et al., 2014; Toulotte, Fabre, Dangremont, Lensel, & Thévenon, 2003; Wiloth, Werner, Lemke, Bauer, & Hauer, 2018; Yoon et al., 2013 vs. Suttanon et al., 2013) and in thirteen of sixteen RCT applying balance scales reaching medium to large effect sizes (Arcoverde et al., 2014; Christofoletti et al., 2008; Dawson, Judge, & Gerhart, 2019; Francese, Sorrell, & Butler, 1997; Hauer et al., 2012; Hauer et al., 2017; M.-J. Kim et al., 2016; Kovács, Sztruhár Jónásné, Karóczi, Korpos, & Gondos, 2013; Padala et al., 2017; Santana-Sosa, Barriopedro, López-Mojares, Pérez, & Lucia, 2008; Telenius, Engedal, & Bergland, 2015a; Toots et al., 2017; Yoon et al., 2013 vs. Burgener et al., 2008; Lam, Liao, Kwok, & Pang, 2018; Padala et al., 2012). In contrast,

only one RCT examining dynamic balance reports statistically significant improvements (Bossers et al., 2015), while two others do not (Suttanon et al., 2013; Wesson et al., 2013).

In line with findings observed in individual RCT, six of seven recent reviews report positive effects on balance (Blankevoort et al., 2010; Brett et al., 2016; Hernández et al., 2015; Lam, Huang et al., 2018; Potter et al., 2011; Suttanon et al., 2010 vs. Littbrand, Stenvall, & Rosendahl, 2011). More detailed, their estimations range from potential improvements (Hernández et al., 2015) through limited evidence (Potter et al., 2011) to strong evidence (Lam, Huang et al., 2018), while pooled effect sizes of 1.08, [0.31 to 3.79] (large effect; Blankevoort et al., 2010) and effect sizes ranging between 0.07, [-0.62 to 0.76] (no effect) and 3.29, [2.17 to 4.41] (large effect; Suttanon et al., 2010) are reported. With predominately moderate to large effect sizes, there are indications that the effects of physical activity on balance and mobility are greater than on other motor domains (Suttanon et al., 2010).

Based on observations in RCT and recent reviews, it can be concluded that physical activity has a positive effect on balance. This especially applies for balance assessed with balance scales, but also for static balance and postural sway. Conclusions on the effects on dynamic balance cannot be drawn, as only a few RCT with inconsistent results are available and recent reviews do not distinguish between different balance abilities.

3.1.2 Mobility

The effectiveness of physical activity on mobility is assessed in 31 RCT. A detailed consideration of spatiotemporal gait parameters, often assigned to mobility, is provided in Chapter 3.2. Accordingly, appropriate results are not more closely considered in this section but are taken into account for overall conclusions. Regarding get up and go tasks, results are inconsistent. Nearly half of identified RCT observe statistically significant improvements with small to large effect sizes (Arcoverde et al., 2014; Cancela, Ayán, Varela, & Seijo, 2016; Hauer et al., 2012; Kampragkou et al., 2017; Kovács et al., 2013; Santana-Sosa et al., 2008; Toulotte et al., 2003; Vreugdenhil et al., 2012), while the remaining ones report no effects (Aguiar, Monteiro, Feres, Gomes, & Melo, 2014; Bossers et al., 2015; Christofoletti et al., 2008; Lam, Liao et al., 2018; Netz et al., 2007; Padala et al., 2012; Rolland et al., 2007; Sobol et al., 2016; Suttanon et al., 2013; Yoon et al., 2013). Similarly, inconsistent results are reported for spatiotemporal

gait parameters (see Chapter 3.2) and mobility index scores (Roach, Tappen, Kirk-Sanchez, Williams, & Loewenstein, 2011 vs. Pomeroy et al., 1999; Roach et al., 2011; Schwenk, Dutzi et al., 2014).

Compared to individual RCT, findings of recent reviews are more consistent, with six of seven reviews showing improvements in mobility (Blankevoort et al., 2010; Brett et al., 2016; Lam, Huang et al., 2018; Pitkälä, Savikko et al., 2013; Potter et al., 2011; Suttanon et al., 2010) and one reporting inconsistent results (Littbrand et al., 2011). Based on these findings, recent reviews conclude on limited (Littbrand et al., 2011), moderate (Pitkälä, Savikko et al., 2013), and strong evidence (Lam, Huang et al., 2018). Small pooled effect sizes (0.28, [-0.25 to 2.37]; Blankevoort et al., 2010) and no (0.00, [-0.36 to 0.36]) to large effects (1.85, [0.97 to 2.73]; Suttanon et al., 2010), respectively, are reported. With predominately moderate to large effect sizes, effects of physical activity on balance and mobility are greater than on other motor domains (Suttanon et al., 2010).

Overall, inconsistency in RCT does not allow drawing clear conclusions on the effectiveness of physical activity on mobility in IWD. Nevertheless, recent reviews clearly support effectiveness and several previous RCT also suggest potential effectiveness.

3.1.3 Strength

Aiming to examine the effects of physical activity on strength, 16 RCT apply different strength assessments. For simple sit-to-stand (STS) assessments, eight RCT report statistically significant improvements with small to large effect sizes (Arcoverde et al., 2014; Bossers et al., 2015; Dawson et al., 2019; Hauer et al., 2012; Hauer et al., 2017; Santana-Sosa et al., 2008; Schwenk, Dutzi et al., 2014; Vreugdenhil et al., 2012), while five observe no effects (Lam, Liao et al., 2018; Netz et al., 2007; Sobol et al., 2016; Suttanon et al., 2013; Telenius et al., 2015a). Comparably, qualitative ratings of STS performance show statistically significant improvements with large effect sizes (Werner et al., 2017), while results of technical examinations are inconsistent (Hauer et al., 2017; Werner et al., 2017 vs. Suttanon et al., 2013). Moreover, all RCT assessing lower-limb strength with dynamometers or fitness machines observe statistically significant improvements with medium to large effect sizes (Bossers et al., 2015; Hauer et al., 2012; Schwenk, Dutzi et al., 2014). Additionally, statistically significant effects are shown for other strength assessments (Francese et al., 1997; Hauer et al., 2012; Santana-Sosa et al., 2008), which however are just applied in one RCT in each case.

In contrast, results concerning handgrip strength are inconsistent with one RCT reporting statistically significant improvements (M.-J. Kim et al., 2016) and two showing no effects (Hauer et al., 2012; Schwenk, Dutzi et al., 2014).

Similar to previous RCT, five of five reviews report positive effects of physical activity on strength (Blankevoort et al., 2010; Hernández et al., 2015; Lam, Huang et al., 2018; Potter et al., 2011; Suttanon et al., 2010). Most focus on lower-limb strength, but one also takes upper-limb strength into account (Suttanon et al., 2010). Despite these consistent findings, derived conclusions are different and range from potential improvements (Hernández et al., 2015) to strong evidence (Lam, Huang et al., 2018). For lower-limb strength, a pooled effect size of 0.85, [-0.04 to 3.14] (large effect) is reported (Blankevoort et al., 2010).

Apart from handgrip strength, the majority of RCT and all recent reviews report positive effects of physical activity on different strength outcomes. Accordingly, it can be concluded that physical activity is effective in improving lower limb strength in IWD.

3.1.4 Endurance

Endurance is assessed in eleven RCT. Those applying 6-minute walk test (6min WT) or 2-min step tests mainly report statistically significant improvements (Bossers et al., 2015; Miu et al., 2008; Santana-Sosa et al., 2008; Tappen, Roach, Applegate, & Stowell, 2000; Venturelli, Scarsini, & Schena, 2011 vs. Roach et al., 2011), while RCT utilizing other versions of simple walk tests (WT) observe no effects (Cott, Dawson, Sidani, & Wells, 2002; Pomeroy et al., 1999; Sobol et al., 2016). Moreover, a 3-speed walking test (Pedrinolla et al., 2018) and cycle ergometer tests (M.-J. Kim et al., 2016; Sobol et al., 2016) show statistically significant improvements of endurance.

Comparably to RCT, three of three reviews observe positive effects of physical activity on endurance (Blankevoort et al., 2010; Hernández et al., 2015; Lam, Huang et al., 2018) and one reports large effect sizes (1.08, [0.31 to 3.79]; Blankevoort et al., 2010). Again, derived conclusions range from potential improvements (Hernández et al., 2015) to strong evidence (Lam, Huang et al., 2018).

With some exceptions, previous RCT and recent reviews observe statistically significant improvements in endurance outcomes. Accordingly, it can be concluded that physical activity has positive effects on endurance.

3.1.5 Flexibility

The effectiveness of physical activity on flexibility is rarely assessed in previous RCT and recent reviews. Two RCT (Santana-Sosa et al., 2008; Toulotte et al., 2003) and four of four reviews (Hernández et al., 2015; Lam, Huang et al., 2018; Potter et al., 2011; Suttanon et al., 2010) observe statistically significant improvements with large effect sizes for lower- and upper-limb flexibility. Based on these few RCT and recent reviews no reliable conclusions are possible. In line with this, the evidence seems to be small, as reviews conclude on potential improvements (Hernández et al., 2015) and weak evidence (Lam, Huang et al., 2018). The reported positive effects nevertheless indicate that physical activity may be effective in improving flexibility in IWD.

3.1.6 Functional performance

Besides taking into account single motor domains, previous RCT also consider functional performance combing different motor domains or focusing on performance-based ADL. One RCT observes statistically significant improvements in the Short Physical Performance Battery (SPPB) (Hauer et al., 2017), while two found no effects (Pitkälä, Pöysti et al., 2013; Souto Barreto et al., 2017). With respect to performance-based ADL, one RCT shows statistically significant improvements (Bossers et al., 2016) and another one reports no effects (Henskens, Nauta, Drost, & Scherder, 2018). For all other functional performance measures no effects are observed (Steinberg, Leoutsakos, Podewils, & Lyketsos, 2009; Suttanon et al., 2013; Wesson et al., 2013). Comparable outcomes are not considered in recent reviews. Based on these inconsistent results, no clear conclusions towards the effectiveness of physical activity on functional performance in IWD are possible.

3.1.7 Effectiveness of physical activity on motor performance in relation to applied physical activity interventions

Based on the expected influences of applied interventions, the effectiveness of physical activity on motor performance is considered in relation to the type and duration of different exercise programs utilized in previous RCT supplemented by appropriate conclusions of recent reviews. Therefore, Table 6 groups previous RCT in those reporting positive effects and those showing no effects for different interventions and motor outcomes.

Table 6. Effectiveness of physical activity on different motor domains in relation to applied interventions grouping previous randomized controlled trials in those reporting positive effects and those showing no effects

	Number of studies showing positive/no effects of multimodal exercise on motor performance	**Number of studies showing positive/no effects of aerobic exercise on motor performance**	**Number of studies showing positive/no effects of other exercises on motor performance**
Balance	*Positive effects*: 12 (Bossers et al., 2015; Christofoletti et al., 2008; Dawson et al., 2019; Hauer et al., 2012; Hauer et al., 2017; Kovács et al., 2013; Santana-Sosa et al., 2008; Schwenk, Dutzi et al., 2014; Telenius et al., 2015a; Toots et al., 2016; Toulotte et al., 2003; Vreugdenhil et al., 2012) *No effects*: 4 (Netz et al., 2007; Rolland et al., 2007; Suttanon et al., 2013; Wesson et al., 2013)	*Positive effects*: 5 (Arcoverde et al., 2014; Kampragkou et al., 2017; M.-J. Kim et al., 2016; Miu et al., 2008; Yoon et al., 2013) *No effects*: 1 (Bossers et al., 2015)	*Positive effects*: 1 (Wiloth et al., 2018) → Balance training *Positive effects*: 1 (Padala et al., 2017) *No effects*: 1 (Padala et al., 2012) → Wii-Fit intervention *Inconsistent effects*: 1 (Burgener et al., 2008) → Taiji *No effects*: 1 (Lam, Liao et al., 2018) → Whole body vibration training
Mobility	*Positive effects*: 11 (Dawson et al., 2019; Hauer et al., 2012; Kemoun et al., 2010; Kovács et al., 2013; Roach et al., 2011; Rolland et al., 2007; Santana-Sosa et al., 2008; Schwenk, Zieschang et al., 2014; Schwenk, Zieschang, Oster, & Hauer, 2010; Toulotte et al., 2003; Vreugdenhil et al., 2012) *No effects*: 12 (Aguiar et al., 2014; Bossers et al., 2015; Christofoletti et al., 2008; Hauer et al., 2017; Netz et al., 2007; Pedrinolla et al., 2018; Schwenk, Dutzi et al., 2014; Souto Barreto et al., 2017; Steinberg et al., 2009; Suttanon et al., 2013; Telenius et al., 2015a; Toots et al., 2017)	*Positive effects*: 3 (Arcoverde et al., 2014; Cancela et al., 2016; Kampragkou et al., 2017) *No effects*: 3 (Bossers et al., 2015; Sobol et al., 2016; Yoon et al., 2013)	*No effects*: 1 (Padala et al., 2012) → Wii-Fit intervention *No effects*: 1 (Lam, Liao et al., 2018) → Whole body vibration training *No effects*: 1 (Pomeroy et al., 1999) → Two-weeks individual physiotherapy
Strength	*Positive effects*: 8 (Bossers et al., 2015; Dawson et al., 2019; Hauer et al., 2012; Hauer et al., 2017; Santana-Sosa et al., 2008; Schwenk, Dutzi et al., 2014; Vreugdenhil et al., 2012; Werner et al., 2017) *No effects*: 3 (Netz et al., 2007; Suttanon et al., 2013; Telenius et al., 2015a)	*Positive effects*: 2 (Arcoverde et al., 2014; M.-J. Kim et al., 2016) *No Effects*: 2 (Bossers et al., 2015; Sobol et al., 2016)	*No effects*: 1 (Lam, Liao et al., 2018) → Whole body vibration training

Endurance	*Positive effects*: Bossers et al., 2015; Pedrinolla et al., 2018; Santana-Sosa et al., 2008; Venturelli et al., 2011 *No effects*: 1 (Roach et al., 2011)	*Positive effects*: 5 (Bossers et al., 2015; M.-J. Kim et al., 2016; Miu et al., 2008; Sobol et al., 2016; Tappen et al., 2000) *No effects*: 1 (Cott et al., 2002)	*No effects*: 1 (Pomeroy et al., 1999) → Two-weeks individual physiotherapy
Flexibility	*Positive effects*: 2 (Santana-Sosa et al., 2008; Toulotte et al., 2003)	Not applied	Not applied
Functional performance	*Positive effects*: 2 (Bossers et al., 2016; Hauer et al., 2017) *No effects*: 6 (Henskens et al., 2018; Pitkälä, Pöysti et al., 2013; Souto Barreto et al., 2017; Steinberg et al., 2009; Suttanon et al., 2013; Wesson et al., 2013)	*Positive effects*: 1 (Bossers et al., 2016, smaller than the effects of multimodal exercise)	Not applied

Comparing the number of RCT reporting positive effects with those showing no effects, findings for different physical activity interventions often do not allow clear assumptions. Overall, most RCT apply multimodal exercise including at least two components of balance, strength, endurance, walking, functional training, coordination, and flexibility. Interventions focusing on other exercises are rarely investigated which hampers drawing conclusions on the type of physical activity. Multimodal exercise seems to be effective for all motor domains, with some restrictions for mobility. Similarly, most RCT applying aerobic exercise report positive effects for balance, endurance, and functional performance, while only potential effectiveness for mobility and strength can be assumed, due to inconsistent findings. In contrast, other types of exercise are less effective in enhancing different motor outcomes, with exceptions for balance training improving balance. However, such interventions are rarely applied and some types of interventions such as single-component strength training are not included in identified RCT. Furthermore, conclusions regarding the duration of physical activity can only be drawn for mobility, as no clear associations are observed for other motor domains. Considering mobility, physical activity interventions need to comprise at least twelve weeks, as no positive effects are reported in shorter interventions (Bossers et al., 2015; Hauer et al., 2017; Lam, Liao et al., 2018; Padala et al., 2012; Pomeroy et al., 1999).

Findings towards applied interventions and effects observed in previous RCT are in line with conclusions of recent reviews. With respect to the type of physical activity, recent reviews support two hypotheses:

1. Multimodal interventions (e.g. a combination of endurance, strength, and balance) seem to be more effective than single-component ones (Blankevoort et al., 2010; Brett et al., 2016; Hernández et al., 2015; Lam, Huang et al., 2018; Pitkälä, Savikko et al., 2013)
2. Interventions are beneficial when they are targeted on and include components related to the outcome measure (Blankevoort et al., 2010; Brett et al., 2016; Lam, Huang et al., 2018; Littbrand et al., 2011)

Additionally, appropriate intensity and duration seem to be important. Pitkälä, Savikko et al. (2013) conclude that both, intensive and long-term physical activity interventions may improve motor performance. In line with this, two other reviews observe that interventions with higher intensity or larger training volumes are more effective than less intensive or shorter ones (Blankevoort et al., 2010; Potter et al., 2011). Accordingly,

interventions should include regular intensive physical activity of around 45-60min, two to three times per week for at least 12 to 16 weeks to provide sufficient intensity/duration allowing changes in motor performance (Blankevoort et al., 2010; Brett et al., 2016; Lam, Huang et al., 2018; Pitkälä, Savikko et al., 2013).

3.2 Effectiveness of physical activity on gait performance in individuals with dementia

The current state of research on the effectiveness of physical activity on gait performance in IWD considers the findings of 32 previous studies and six recent reviews. This chapter summarizes related findings for single and dual task conditions and additionally focuses on possible influences, such as study designs, methods to assess spatiotemporal gait parameters, and interventions.

3.2.1 Spatiotemporal gait parameters (single task condition)

Walking speed and time to walk a certain distance, respectively, are most frequently assessed in previous studies. Eleven of thirty studies observe statistically significant improvements of walking speed/time to walk a certain distance in intervention group (IG) or time*group interaction effects in favor for IG (Ahn & Kim, 2015; Aman & Thomas, 2009; Bossers, Scherder et al., 2014; Hauer et al., 2012; Kemoun et al., 2010; J.-S. Kim, Kang, Moon, & Oh, 2017; Manckoundia, Taroux, Kubicki, & Mourey, 2014; Perrochon, Tchalla, Bonis, Perucaud, & Mandigout, 2015; Rolland et al., 2007; Schwenk, Zieschang et al., 2014; Toulotte et al., 2003). All other studies report no effects (Bossers et al., 2015; Cadore et al., 2014; Dawson et al., 2019; Hageman & Thomas, 2002; Hauer et al., 2017; Junge, Knudsen, & Kristensen, 2018; Kuiack, Campbell, & Evans, 2004; McCaffrey, Park, Newman, & Hagen, 2014; Pedrinolla et al., 2018; Ries, Hutson, Maralit, & Brown, 2015; Schwenk, Dutzi et al., 2014; Sobol et al., 2016; Souto Barreto et al., 2017; Steinberg et al., 2009; Suttanon et al., 2013; Tay, Lim, Chan, Ali, & Chong, 2016; Telenius et al., 2015a; Thomas & Hageman, 2003; Toots et al., 2017). Thus, results on the effectiveness of physical activity on walking speed in IWD are inconsistent.

In contrast to inconsistent findings of previous studies, six of six recent reviews consistently observe statistically significant effects in majority of included studies and thus conclude that physical activity is effective in improving walking speed in IWD (Blankevoort et al., 2010; Brett et al., 2016; Lam, Huang et al., 2018; Littbrand et al.,

2011; Potter et al., 2011; Suttanon et al., 2010). However, they differently rate the quality of evidence varying between some (Littbrand et al., 2011) and strong evidence of moderate quality (Lam, Huang et al., 2018). More detailed, they report statistically significant improvements with weighted mean differences of 0.06 m/s, [0.01 to 0.1] and 0.13 m/s [0.03 to 0.24], respectively (Lam, Huang et al., 2018; Potter et al., 2011). Related effect sizes are small (d=0.29, [-0.11 to 0.50]; Blankevoort et al., 2010) or range between small (standardized mean difference=0.23, [-0.13 to 0.60]) and medium effects (standardized mean difference=0.55, [0.10 to 1.01]; Suttanon et al., 2010)[4].

With respect to further spatiotemporal gait parameters, previous studies predominately show statistically significant improvements in IG or time*group interaction effects in favor for IG for stride length (Coelho et al., 2013; Kemoun et al., 2010; J.-S. Kim et al., 2017; Orcioli-Silva et al., 2018; Perrochon et al., 2015; Schwenk, Zieschang et al., 2014 vs. Pedrinolla et al., 2018; Schwenk, Dutzi et al., 2014), stride time (Orcioli-Silva et al., 2018; Schwenk, Zieschang et al., 2014), step time (J.-S. Kim et al., 2017), double support time (Kemoun et al., 2010), and stride frequency (Perrochon et al., 2015). With few exceptions, no effects are reported for step length (Bossers et al., 2015; Hageman & Thomas, 2002; Pedrinolla et al., 2018; Suttanon et al., 2013 vs. Hauer et al., 2012), step width (Schwenk, Zieschang et al., 2014; Suttanon et al., 2013), and percent of single support (Pedrinolla et al., 2018), while results are inconsistent for stride speed (Orcioli-Silva et al., 2018 vs. Coelho et al., 2013), percent of double support (Schwenk, Zieschang et al., 2014 vs. Orcioli-Silva et al., 2018; Pedrinolla et al., 2018), and cadence (Hauer et al., 2012; J.-S. Kim et al., 2017; Orcioli-Silva et al., 2018; Schwenk, Zieschang et al., 2014 vs. Coelho et al., 2013; Schwenk, Dutzi et al., 2014). These inconsistent findings and the small number of available studies (predominately one to two studies for each parameter, maximal eight studies) compromise drawing meaningful conclusions on the effectiveness of physical activity on further spatiotemporal gait parameters in IWD.

Recent reviews rarely consider further spatiotemporal gait parameters. Herein, statistically significant improvements are reported for stride/step length and double support (Brett et al., 2016; Lam, Huang et al., 2018). While no more detailed information is available for stride length and double support, statistically significant effects resulting

[4] These effect sizes also refer to walking speed assessed during 2min WT. Recent reviews do not strictly distinguish between walking performance of different assessments. This thesis only considers spatiotemporal gait parameters of short distance walk tests and instrumented gait analyses. However, no clear differentiation was possible in this case.

in an increased step length of 5 cm, [2 to 8] are revealed and strong evidence of high quality is concluded (Lam, Huang et al., 2018).

Overall, previous studies and recent reviews support the effectiveness of physical activity on gait performance for stride length and double support, and with some restrictions for walking speed and step length. Moreover, indications for improving stride/step time and stride frequency/cadence can be suggested based on findings of previous studies.

3.2.2 Spatiotemporal gait parameters (dual task condition)

Besides assessing spatiotemporal gait parameters in single task conditions, six previous studies additionally examine the effectiveness of physical activity on gait performance in dual task conditions (naming animals, counting backward). They observe positive effects for stride time (Orcioli-Silva et al., 2018) and no effects for double support (Orcioli-Silva et al., 2018), while findings on walking speed, stride length, stride speed, and cadence are inconsistent (Cadore et al., 2015; Coelho et al., 2013; Junge et al., 2018; Orcioli-Silva et al., 2018; Sobol et al., 2016; Tay et al., 2016). Furthermore, Schwenk et al. (2010) also investigate gait performance in dual task conditions but only report dual task costs and no spatiotemporal gait parameters. While they observe statistically significant time*group interaction effects in favor of IG for almost all parameters with serial three backward calculation, no effects are shown for less complex serial two forward calculation.

The effectiveness of physical activity on gait performance in dual task conditions in IWD is rarely examined in recent reviews. Only one review refers to this outcome and found weak evidence of low quality against the effectiveness of non-specific exercise for improving dual task performance (Lam, Huang et al., 2018). These findings are not specifically related to spatiotemporal gait parameters. Two of three studies including IWD apply the cognitive Timed Up & Go Test (TUG), while the third considers dual task cost but no spatiotemporal gait parameters itself.

Summarizing these findings of previous studies and recent reviews, no clear conclusion on the effectiveness of physical activity on gait performance in IWD can be drawn for dual task conditions. Accordingly, the effectiveness of physical activity on spatiotemporal gait parameters during dual task conditions is questionable.

3.2.3 Factors potentially impacting findings on spatiotemporal gait parameters

Comparing studies investigating the effectiveness of physical activity on spatiotemporal gait parameters reveals large heterogeneity with respect to study designs, methods to assess spatiotemporal gait parameters, and interventions (see Table 7). These factors may affect findings on the effectiveness of physical activity and thus need to be considered.

Table 7. Heterogeneity of previous studies assessing the effectiveness of physical activity on spatiotemporal gait parameters considering study design, assessment methods, and applied interventions

		Studies
Study design	(Cluster) randomized controlled trials	Bossers et al., 2015; Dawson et al., 2019; Hauer et al., 2012; Hauer et al., 2017; Kemoun et al., 2010; Pedrinolla et al., 2018; Rolland et al., 2007; Schwenk, Dutzi et al., 2014; Schwenk, Zieschang et al., 2014; Sobol et al., 2016; Souto Barreto et al., 2017; Steinberg et al., 2009; Suttanon et al., 2013; Telenius et al., 2015a; Toots et al., 2017; Toulotte et al., 2003
	Controlled trials	Bossers, Scherder et al., 2014; Coelho et al., 2013; Manckoundia et al., 2014; Orcioli-Silva et al., 2018
	Uncontrolled trials	Ahn & Kim, 2015; Aman & Thomas, 2009; Cadore et al., 2014; Hageman & Thomas, 2002; Junge et al., 2018; J.-S. Kim et al., 2017; Kuiack et al., 2004; McCaffrey et al., 2014; Perrochon et al., 2015; Ries et al., 2015; Tay et al., 2016; Thomas & Hageman, 2003
Assessment methods	Short distance walk tests	*4 m*: Souto Barreto et al., 2017; Toots et al., 2017 *5 m*: Cadore et al., 2014 *6 m*: Aman & Thomas, 2009; Bossers et al., 2015; Bossers, Scherder et al., 2014; Hageman & Thomas, 2002; Kuiack et al., 2004; McCaffrey et al., 2014; Rolland et al., 2007; Tay et al., 2016; Telenius et al., 2015a; Thomas & Hageman, 2003 *8 m*: Ahn & Kim, 2015 *10 m*: Junge et al., 2018; Manckoundia et al., 2014; Sobol et al., 2016; Toulotte et al., 2003 *8 ft*: Dawson et al., 2019; Hauer et al., 2012; Hauer et al., 2017; Steinberg et al., 2009
	Instrumented gait analysis systems	*GAITRite*: J.-S. Kim et al., 2017; Pedrinolla et al., 2018; Ries et al., 2015; Schwenk, Dutzi et al., 2014; Schwenk, Zieschang et al., 2014 *Kinematic video analysis*: Coelho et al., 2013; Orcioli-Silva et al., 2018 *Locométrix triaxial accelerometer*: Perrochon et al., 2015 *NeuroCom Balance Master*: Suttanon et al., 2013 *Bessou locometer*: Kemoun et al., 2010
Applied interventions	Multimodal exercise	Aman & Thomas, 2009; Bossers et al., 2015; Bossers, Scherder et al., 2014; Cadore et al., 2014; Coelho et al., 2013; Dawson et al., 2019; Hauer et al., 2012; Hauer et al., 2017; Junge et al., 2018; Kemoun et al., 2010; Manckoundia et al., 2014; Orcioli-Silva et al., 2018; Pedrinolla et al., 2018; Perrochon et al., 2015; Rolland et al., 2007; Schwenk, Dutzi et al., 2014; Schwenk, Zieschang et al., 2014; Souto Barreto et al., 2017; Steinberg et al., 2009; Suttanon et al., 2013; Tay et al., 2016; Telenius et al., 2015a; Toots et al., 2017; Toulotte et al., 2003

Single-component exercise	*Pelvic tilt exercises*: J.-S. Kim et al., 2017 *Balance training*: Ries et al., 2015 *Resistance training*: Ahn & Kim, 2015; Hageman & Thomas, 2002; Kuiack et al., 2004; Thomas & Hageman, 2003 *Aerobic exercise*: Sobol et al., 2016 *Chair yoga*: McCaffrey et al., 2014

In previous studies, findings between different study designs are comparable. Nevertheless, most evidence generally can be derived from RCT. Considering recent reviews, all assess the methodological quality of included studies and one performs sensitivity analyses for walking speed only including studies with high methodological quality still resulting in statistically significant but slightly smaller improvements of 0.08 m/s, [0.01 to 0.15] (Lam, Huang et al., 2018). Accordingly, RCT of high methodological quality are most meaningful when concluding on evidence.

With respect to assessment methods, instrumented gait analysis systems enable detailed examinations of gait performance, while simple short distance WT do not allow to measure various spatiotemporal gait parameters. Accordingly, the influence of the assessment method can only be examined for walking speed, which is easily assessable by both methods. Herein, less than one third of those studies applying simple short distance WT report statistically significant effects. In contrast, one half of those using instrumented gait analysis systems support the effectiveness of physical activity on walking speed. As recent reviews do not include information on influences of assessment methods, no further conclusions can be drawn in this context. This is also related to the fact that all reviews focus on short distance WT and only three consider studies utilizing instrumented gait analysis. Based on observations in previous studies, instrumented gait analysis systems seem to be more appropriate in examining the effectiveness of physical activity on spatiotemporal gait parameters. Compared to simple short distance WT, instrument gait analysis systems allow determining various gait parameters and may be more sensitive for small intervention-induced changes in spatiotemporal gait parameters.

Comparing findings of previous studies with different types of interventions shows similar proportions of studies with and without statistically significant effects. This observation does not support the superiority of one type of physical activity. However, it is important to take into account that often only individual studies investigate these interventions, which may affect conclusions. In contrast, recent reviews consistently observe lager improvements of multimodal exercise compared to progressive resistance

training alone (Blankevoort et al., 2010; Lam, Huang et al., 2018). Considering single previous studies combining motor and cognitive exercise, those performing both task simultaneously (Coelho et al., 2013; Orcioli-Silva et al., 2018) rather contribute to improvements in spatiotemporal gait parameters than those without coincidental interventions of motor and cognitive exercise (Cadore et al., 2014; Junge et al., 2018; Tay et al., 2016). Focusing on the duration and/or intensity, recent reviews particularly find improvements for high-intensity interventions (Potter et al., 2011) or interventions comprising physical activity of 45 to 120 min, two to three times per week for at least 15 weeks (Lam, Huang et al., 2018). Overall, especially interventions applying multimodal exercise of sufficient intensity and duration may be effective. Additionally, promising indications for combined motor and cognitive exercises are available.

3.3 Effectiveness of physical activity on cognitive performance in individuals with dementia

This chapter comprises a short summary of the current state of research on the effectiveness of physical activity on cognitive performance in IWD, which is based on nine recent reviews. These reviews focus on cognition in general or global cognition and rarely consider cognitive subdomains.

3.3.1 Global cognition

Six of nine reviews observe statistically significant improvements in global cognition or reduced decline in the majority of studies included and derive evidence for effectiveness of physical activity (Brett et al., 2016; Du et al., 2018; Farina et al., 2014; Groot et al., 2016; Hernández et al., 2015; Karssemeijer et al., 2017). However, this relationship is not observed in all studies, and recent reviews refer to the risk of bias, low methodological quality, and methodological heterogeneity. Thus, conclusions are limited and should be considered with caution (Brett et al., 2016; Du et al., 2018; Farina et al., 2014). Meta-analyses in recent reviews reveal standardized mean differences of 0.75, [0.32 to 1.17] (Farina et al., 2014), 0.42, [0.23 to 0.62] ($p<0.01$; Groot et al., 2016), and 0.36, [0.12 to 0.60] ($p<0.00$; Karssemeijer et al., 2017), respectively, indicating small to medium effects. In contrast, three reviews report no clear effects and/or evidence of physical activity on global cognition (Forbes, Forbes, Blake, Thiessen, & Forbes, 2015; Littbrand et al., 2011; Öhman, Savikko, Strandberg, & Pitkälä, 2014). All three reviews refer to low methodological quality as well as inadequate interventions

of insufficient intensity and conclude that the effectiveness of physical activity on cognition in IWD remains unclear, possibly ranging between minimal harms and substantial benefit (Forbes et al., 2015; Littbrand et al., 2011; Öhman et al., 2014).

3.3.2 Cognitive subdomains

Recent reviews rarely assess the effectiveness of physical activity on specific cognitive subdomains. In this context, Farina et al. (2014) indicate an insufficient number of studies to explore the effects on cognitive subdomains. Inconsistent findings in recent reviews reflect this insufficient research. While Hernández et al. (2015) conclude that exercise may improve cognitive subdomains such as sustained attention, visual memory, and frontal cognitive function, Karssemeijer et al. (2017) observe no effects in domain-specific analysis for executive function/attention (standardized mean difference=0.38, [-0.21 to 0.97] ,$p>0.05$) and memory (standardized mean difference=0.02, [-0.35 to 0.39], $p>0.05$). In between, Öhman et al. (2014), observe no effects on cognitive subdomains in studies of moderate quality but report some effects on executive function in studies of poorer quality. Moreover, they find no effects on attention, delayed recall, working memory, declarative memory.

3.3.3 Effectiveness of physical activity on cognitive performance in relation to applied physical activity interventions

When discussing the effectiveness of physical activity on cognition, it is important to conclude on the characteristics of effective interventions. However, appropriate findings are rare in recent reviews. Farina et al. (2014) refer to the heterogeneity of study designs, which imped conclusions on optimum interventions. Other reviews suggest interventions consisting of several components and/or including aerobic tasks (Brett et al., 2016; Groot et al., 2016). Referring to this, positive effects are observed for multimodal (standardized mean difference=0.59, [0.32 to 0.86], $p<0.01$) and aerobic interventions (standardized mean difference=0.41, [0.05 to 0.76], $p<0.05$), but not for single-component non-aerobic interventions (standardized mean difference=-0.10, [-0.38 to 0.19], $p=0.51$; Groot et al., 2016). However, this does not include interventions only compromising walking exercise at a self-selected pace (Littbrand et al., 2011). Furthermore, positive effects are also observed for interventions combining motor and cognitive tasks (Karssemeijer et al., 2017).

3.4 Limitations in previous studies examining the effectiveness of physical activity in individuals with dementia

When drawing conclusions or deriving evidence, recent reviews consistently indicate limitations or risk of bias of previous studies (Blankevoort et al., 2010; Brett et al., 2016; Du et al., 2018; Farina et al., 2014; Forbes et al., 2015; Groot et al., 2016; Hernández et al., 2015; Karssemeijer et al., 2017; Lam, Huang et al., 2018; Littbrand et al., 2011; Öhman et al., 2014; Pitkälä, Savikko et al., 2013; Potter et al., 2011; Suttanon et al., 2010). Generally, several previous studies are of poor quality and suffer from methodological problems (Blankevoort et al., 2010; Öhman et al., 2014; Pitkälä, Savikko et al., 2013; Suttanon et al., 2010). A short summary of substantial limitations is given thereafter.

Several studies, do not adhere to recommendations of Consolidated Standards of Reporting Trials (CONSORT) statements (Hauer, Becker, Lindemann, & Beyer, 2006). For example, information on sample size calculation, randomization, details on interventions, or dropouts is imprecise or missing (Blankevoort et al., 2010; Brett et al., 2016; Forbes et al., 2015; Littbrand et al., 2011; Öhman et al., 2014; Pitkälä, Savikko et al., 2013; Potter et al., 2011; Suttanon et al., 2010). Accordingly, applied methods are not comprehensible, which impedes conclusions on actual methodological quality. Moreover, recent reviews frequently criticize small sample, as they result in low statistical power (Brett et al., 2016; Du et al., 2018; Öhman et al., 2014; Pitkälä, Savikko et al., 2013; Potter et al., 2011; Suttanon et al., 2010). Other limitations include heterogeneity of participants within studies, which is mostly not considered for assessments and interventions, as well as during analysis (Hauer et al., 2006; Littbrand et al., 2011; Suttanon et al., 2010). At the content level, previous studies frequently show limitations referring to inappropriate, non-standardized, or insensitive motor assessments (Hauer et al., 2006; Suttanon et al., 2010) as well as unspecific and inadequate interventions of short duration (Du et al., 2018; Hauer et al., 2006; Öhman et al., 2014). With respect to data analysis, inadequate statistical methods and infrequent use of intention-to-treat analyses are mentioned (Hauer et al., 2006; Öhman et al., 2014). In this context, only including complete cases in analyses can introduce a potential risk of bias (Öhman et al., 2014; Pitkälä, Savikko et al., 2013).

Besides the limitations of individual studies, heterogeneity of available investigations needs to be considered. For instance, recent reviews indicate heterogeneity related to

methods, participants, interventions, and applied assessments (Blankevoort et al., 2010; Brett et al., 2016; Du et al., 2018; Gonçalves, Cruz, Marques, Demain, & Samuel, 2018; Groot et al., 2016; Karssemeijer et al., 2017; Öhman et al., 2014). These heterogeneities hamper the comparability of previous studies (Blankevoort et al., 2010; Du et al., 2018; Gonçalves et al., 2018; Groot et al., 2016; Karssemeijer et al., 2017; Öhman et al., 2014).

All those limitations and risk of bias potentially influence results (Brett et al., 2016). Findings of studies observing no effects can be caused by limitations or reflect the limited potential of physical activity to positively affect motor and cognitive performance in IWD (Hauer et al., 2006). Together with heterogeneity between studies, limitations prevent drawing firm conclusions and deriving evidence (Blankevoort et al., 2010; Brett et al., 2016; Du et al., 2018; Hauer et al., 2006). Accordingly, recent reviews consistently refer to the urgent need of further high-quality research aiming to support current evidence on the effectiveness of physical activity in IWD (Blankevoort et al., 2010; Brett et al., 2016; Du et al., 2018; Farina et al., 2014; Forbes et al., 2015; Groot et al., 2016; Karssemeijer et al., 2017; Lam, Huang et al., 2018; Littbrand et al., 2011; Öhman et al., 2014; Suttanon et al., 2010).

Summary

Theoretical foundation for establishing a high-quality methodological approach

- Previous studies frequently show limitations related to insufficient reporting, small and heterogeneous samples, inadequate interventions and assessments, as well as inappropriate data analyses. These limitations induce a risk of bias and thus compromise conclusions on evidence. Accordingly, they need to be considered when designing high-quality studies to investigate the effectiveness of physical activity on motor and gait performance in IWD.

Theoretical foundation for performing a high-quality RCT investigating the effectiveness of physical activity on motor performance in IWD

- Findings in previous RCT and recent reviews allow concluding on the effectiveness of physical activity on balance, strength, and endurance. Due to inconsistent findings, some restrictions exist for mobility but overall potential effectiveness can be assumed. In contrast, no clear conclusions are possible for flexibility and functional performance due to insufficient research and inconsistent results. Nevertheless, limitations in previous studies require further high-quality studies. With respect to interventions, especially multimodal exercise of sufficient duration/intensity seems to be effective.

Theoretical foundation for performing a high-quality RCT investigating the effectiveness of physical activity on gait performance, characteristics of responders, and impacts of changes in underlying motor and cognitive performance on changes in gait performance

- Based on previous studies and recent reviews the effectiveness of physical activity on gait performance can be supported for stride length and double support, and with some restrictions for walking speed and step length. Further indications are available for stride/step time and stride frequency/cadence. However, findings of previous studies are inconsistent and some spatiotemporal gait parameters are only rarely considered. There is a need for further research characterized by RCT designs, application of an instrumented gait analysis, and multimodal exercise interventions. Furthermore, research assessing spatiotemporal gait parameters during dual task conditions is rare.
- Recent reviews provide indications for the effectiveness of physical activity on global cognition. Due to the limitations of previous studies, related conclusions should be considered with caution. Information concerning specific cognitive subdomains is rare and does not allow clear assumptions. Again, multimodal exercise seems to be the most effective. Based on these findings, intervention-induced cognitive impacts of physical activity on changes in gait performance are conceivable.

4 Formulation of research questions and hypotheses

4.1 Research gaps and perspectives related to the current state of research

Summarizing the current state of research allows assuming the effectiveness of physical activity on motor and gait performance in IWD. However, evidence cannot be ensured and several research gaps still exist. In order to examine open research questions, it is necessary to address methodological research gaps and related limitations first.

4.1.1 Considerations towards methodological limitations and research gaps

The examination of previous studies and recent reviews identifies various methodological limitations. Among those, especially assessments and interventions seem to have essential influences on determining the effectiveness of physical activity in IWD. Accordingly, the following sections focus on considerations and approaches towards motor assessments and physical activity interventions first and afterward summarize requirements for high-quality studies.

The selection of adequate assessments is identified as an important aspect for detecting intervention-induced effects on motor and gait performance (Blankevoort et al., 2010; Hauer et al., 2006; Suttanon et al., 2010). Accordingly, appropriate, valid, reliable, sensitive, and standardized assessments are required (Blankevoort et al., 2010; Gonçalves et al., 2018; Hauer et al., 2006; Hernández et al., 2015). However, especially in motor assessments, psychometric properties are not thoroughly examined in IWD, which is probably related to the specificity of this population (Blankevoort et al., 2010). Moreover, it is questionable, whether common motor assessments developed for cognitively unimpaired older adults are appropriate for IWD. Considering cognitive impairments and other specific characteristics of IWD, their performance in motor assessments may not be solely determined by motor capacities but potentially influenced by cognitive or emotional status. Associated problems, for instance, refer to not understanding instructions or missing motivation (Hauer et al., 2006). Based on the importance of motor assessments for investigating the effectiveness of physical activity, identifying assessments fulfilling the above-mentioned criteria is essential. However, research on motor assessments in IWD is rare. Accordingly, a comprehensive exami-

nation considering various qualitative and quantitative aspects is required. An important first step is to summarize, analyze, and discuss available findings concerning motor assessments in IWD.

The effectiveness of physical activity certainly dependents on the characteristics of interventions. Previous studies and recent reviews allow some conclusions on the type of exercise, frequency, duration, and intensity (see Chapter 3.1.7, 3.2.3, 3.3.3). Besides general training scientific issues, specific characteristics, capabilities, needs, and preferences of IWD need to be adequately addressed (Forbes et al., 2015). Based on significant differences in motor and cognitive status (Baddeley, Logie, Bressi, Della Sala, & Spinnler, 1986; Manckoundia et al., 2006; Mc Ardle et al., 2017; Perry & Hodges, 1999; Suttanon et al., 2012), effective interventions for IWD may need some modifications compared to those developed for cognitively unimpaired older adults (Suttanon et al., 2010). Appropriate modifications refer to training scientific (e.g. intensity adapted to reduced capacity) and didactic aspects (e.g. communication adapted to language deficits) and may improve feasibility and adherence (Forbes et al., 2015; Suttanon et al., 2010). However, physical activity interventions specifically tailored to capabilities, needs, and further characteristics of IWD are rarely applied in previous studies (Hauer et al., 2006). Another promising approach with regard to interventions is the combination of motor and cognitive tasks within one intervention. Previous studies and recent reviews show the potential effectiveness of physical activity on motor and cognitive performance (see Chapter 3). Moreover, positive effects are also reported in some studies applying cognitive-focused interventions (Thom & Clare, 2011). It is assumed that interventions combining both activities will enhance effectiveness (Thom & Clare, 2011), as observed in cognitively unimpaired older adults (Fabre, Chamari, Mucci, Masse-Biron, & Prefaut, 2002). Additional benefits are particularly shown in cognitive performance (Fabre et al., 2002). Beyond this, combined interventions may also be more effective with respect to gait performance, as it is associated with motor and cognitive performance (Beauchet et al., 2008; Thom & Clare, 2011; Valkanova & Ebmeier, 2017). There are only a few studies combining physical and cognitive activity in IWD. Some include both interventions types but perform each of them separately (e.g. Cadore et al., 2014; Junge et al., 2018; Tay et al., 2016), while studies simultaneously applying motor and cognitive tasks are rare (e.g. Coelho et al., 2013; Orcioli-Silva et al., 2018; Schwenk et al., 2010) and predominately show methodological limitations, which indicates the need for further research.

As mentioned above, high-quality studies are required to be able to answer open research questions (see Chapter 3.4). High-quality studies are characterized by profound designs and methods including accurate reporting, large and homogeneous samples, relevant outcomes measured with valid, reliable, and sensitive assessments, specific interventions of sufficient duration and intensity tailored to IWD, as well as appropriate presentation and statistical analysis of data (Blankevoort et al., 2010; Brett et al., 2016; Du et al., 2018; Farina et al., 2014; Forbes et al., 2015; Groot et al., 2016; Hauer et al., 2006; Hernández et al., 2015; Littbrand et al., 2011; Öhman et al., 2014; Suttanon et al., 2010). Concrete details are given in Table 8.

Table 8. Criteria for high-quality studies indicated in systematic reviews

	Criteria characterizing high-quality studies
Reporting	Compliance with CONSORT statement guidelines and accurate reporting of methodological aspects (randomization, blinding, attrition rates, reasons for dropouts, adherence to interventions), sample characteristics, and details of intervention (detailed descriptions of exercises, duration, frequency, intensity levels, progression; Blankevoort et al., 2010; Forbes et al., 2015; Groot et al., 2016; Hernández et al., 2015; Littbrand et al., 2011; Suttanon et al., 2010)
Sample	Sufficiently large and homogenous samples with respect to severity and etiology of dementia (Hauer et al., 2006; Öhman et al., 2014; Suttanon et al., 2010)
Outcomes and assessments	Relevant outcomes (e.g. motor and cognitive performance, falls, dual task ability, activities of daily living, quality of life) assessed with valid, reliable, and sensitive assessments appropriate for individuals with dementia, allowing a comprehensive evaluation of motor domains and cognitive subdomains (Blankevoort et al., 2010; Farina et al., 2014; Forbes et al., 2015; Hauer et al., 2006; Lam, Liao et al., 2018; Suttanon et al., 2010)
Interventions	Specific interventions of sufficient duration and intensity tailored to individuals with dementia (Brett et al., 2016; Hauer et al., 2006; Öhman et al., 2014; Suttanon et al., 2010)
Statistical analysis	Appropriate and comprehensive presentation of data (e.g. means and standard deviations of baseline and post-intervention performance and changes, comprehensive statistical characteristics) and adequate handling of missing data (e.g. intention-to-treat analysis; Blankevoort et al., 2010; Forbes et al., 2015; Hernández et al., 2015)

CONSORT: Consolidated Standards of Reporting Trials

4.1.2 Research perspectives

Besides indicating the general need for further research investigating the effectiveness of physical activity on motor, cognitive, and gait performance in IWD, recent reviews establish related research perspectives. These comprise but are not limited to moderators and mediators. Exploring possible moderators and mediators is highly relevant and allows conclusions on characteristics of responders and changes in underlying motor and cognitive performance related to observed effects of physical activity. For

example, it is valuable to know whether effects of physical activity vary between participants showing different characteristics and to examine several moderating factors such as sex, age, as well as severity and etiology of dementia (Blankevoort et al., 2010; Littbrand et al., 2011; Suttanon et al., 2010). Moreover, the identification of mediators is important to learn about changes in underlying motor and cognitive performance required to translate positive effects on motor and cognitive subdomains into improvements in functional or cognitive performance on activity/participation-level (e.g. gait performance, ADL; Blankevoort et al., 2010; Farina et al., 2014; Lam, Huang et al., 2018). Knowledge about such possible impacts can help to optimize the effectiveness of physical activity interventions and to enhance individual responses of IWD with different characteristics (Blankevoort et al., 2010; Farina et al., 2014; Suttanon et al., 2010).

4.2 Research questions and hypotheses

A detailed examination of the current state of research emphasizes the potential effectiveness of physical activity in IWD but simultaneously identifies various research gaps. Accordingly, this thesis aims to evaluate the effectiveness of physical activity on motor and gait performance in IWD. Considering the limitations of the current state of research, this aim is pursued in two steps, which are based on the following primary research questions:

Research question A:

How ***high-quality studies*** need to be ***designed*** to enhance evidence concerning the ***effectiveness of physical activity*** on motor and gait performance in IWD? *(RQ_A)*

Research question B:

Is ***physical activity effective*** in reducing the decline of ***motor and gait performance*** in IWD? *(RQ_B)*

With regard to research question A, the current state of research refers to less investigated and inappropriate motor assessments in IWD. Accordingly, secondary research questions A1 and A2 focus on motor assessments for IWD (see Figure 2).

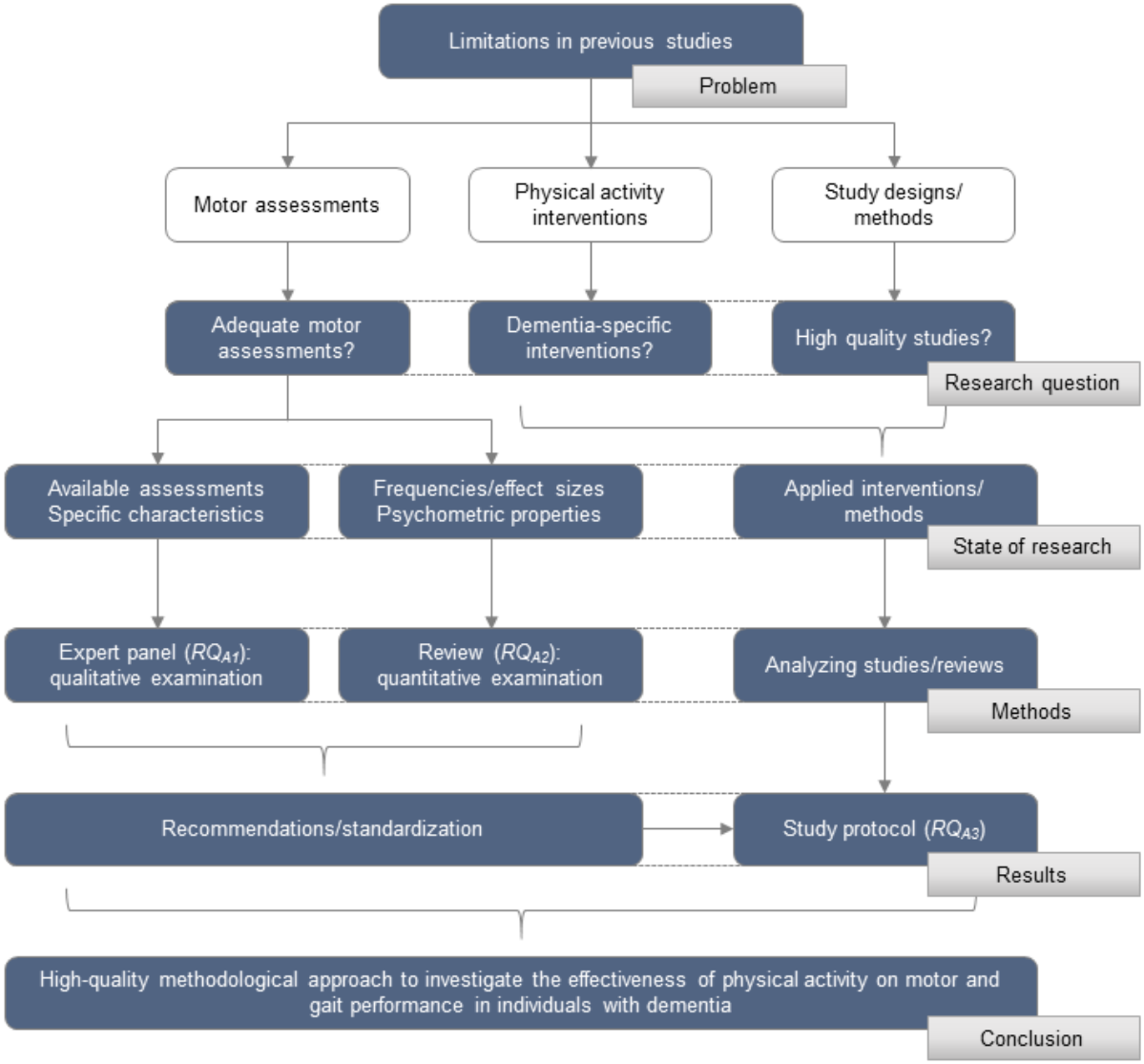


Figure 2. Theoretical considerations and approach to examine primary research question A (RQ: research question).

Research question A1:

Which ***motor assessments*** are ***appropriate*** for IWD based on ***qualitative*** examination?

(RQ_{A1})

Research question A2:

Which ***motor assessments*** can be ***recommended*** for IWD based on ***quantitative*** outcomes, especially psychometric properties?

(RQ_{A2})

Based on observed findings and the examination of previous studies as well as the current state of research, a study design to examine research question B is established (research questions A3). Besides attaching importance to high quality in general and motor assessments adequate for IWD, it is characterized by a dementia-specific MEP, which combines motor and cognitive tasks (see Figure 2).

Research question A3:

How ***high-quality studies*** and ***dementia-specific physical activity interventions*** need to be ***designed*** to investigate the ***effectiveness of physical activity*** on motor and gait performance in IWD?

(RQ_{A3})

Based on the established methodological high-quality approach, an RCT investigating effectiveness of physical activity in IWD contributes to answering research question B. Herein, research question B1 focuses on motor performance, while research questions B2-B4 refer to gait performance.

Research question B1:

Is a dementia-specific MEP combining motor and cognitive tasks effective in ***reducing the decline of motor performance*** in IWD?

(RQ_{B1})

Derived from research question B1, hypothesis 1 is formulated. It is based on promising findings in previous research (see Chapter 3.1, Figure 3).

Hypothesis 1:

A dementia-specific MEP combining motor and cognitive tasks in addition to conventional treatment[5] is more ***effective in reducing the decline of motor performance*** in IWD than conventional treatment alone.

(H_1)

Considering the high prevalence of gait impairments in IWD and their consequence, research questions B2-B4 focus on gait performance. Herein, overall effectiveness, characteristics of responders, and impacts of changes in underlying motor and cognitive performance on changes in gait performance are examined.

Research question B2:

Is a dementia-specific MEP combining motor and cognitive tasks effective in ***reducing the decline of gait performance*** in IWD?

(RQ_{B2})

Derived from research question B2, Hypothesis 2 is formulated. It is based on promising findings of previous research (see Chapter 3.2, Figure 3).

[5] Standard care in care facilities (e.g. medication, care, or therapeutic applications).

Hypothesis 2:

A dementia-specific MEP combining motor and cognitive tasks in addition to conventional treatment[5] is more ***effective in reducing the decline of gait performance*** in IWD than conventional treatment alone.

(H_2)

With respect to heterogeneity in IWD, it is assumed that physical activity interventions do not show the same effect on gait performance in all IWD. Accordingly, research question B3 intends to consider specific characteristics of IWD, while examining positive, non-, and negative responders.

Research question B3:

Regarding gait performance, do ***positive, non-, and negative responders*** of the MEP ***differ in specific characteristics*** closely related to gait performance?

(RQ_{B3})

Derived from research question B3, Hypotheses 3, 4, and 5 are formulated. They are based on relations of gait performance with motor and cognitive performance, as well as the etiology of dementia (see Chapters 2.1.3, 2.2.2, 2.2.3, Figure 3).

Hypothesis 3:

Regarding gait performance, ***positive, non-, and negative responders*** of the MEP differ in the ***motor baseline performance*** (gait, balance, mobility, strength and function of lower limbs, and use of walking aids).

(H_3)

Hypothesis 4:

Regarding gait performance, ***positive, non-, and negative responders*** of the MEP differ in the ***baseline severity of cognitive impairments*** (global cognition, executive function, attention, and working memory).

(H_4)

Hypothesis 5:

Regarding gait performance, ***positive, non-, and negative responders*** of the MEP differ in the ***etiology of dementia*** (AD vs. non-AD).

(H_5)

Considering close associations of gait performance with motor and cognitive performance, impacts of changes in underlying motor and cognitive performance on changes in gait performance will be investigated.

Research question B4:

Which ***changes in underlying motor and cognitive performance*** have an ***impact on changes in gait performance*** in IWD who participated in a dementia-specific MEP?

(RQ_{B4})

Derived from research question B4, Hypotheses 6 and 7 are formulated. They are based on prerequisites and impacts of motor and cognitive performance on changes in gait performance (see Chapter 2.2.2, 2.2.3, Figure 3).

Hypothesis 6:

Changes in underlying motor performance (balance, mobility, strength and function of lower limbs) have an ***impact on changes in gait performance*** in IWD who participated in a dementia-specific MEP.

(H_6)

Hypothesis 7:

Changes in underlying cognitive performance (executive function, attention, and working memory) have an ***impact on changes in gait performance*** in IWD who participated in a dementia-specific MEP.

(H_7)

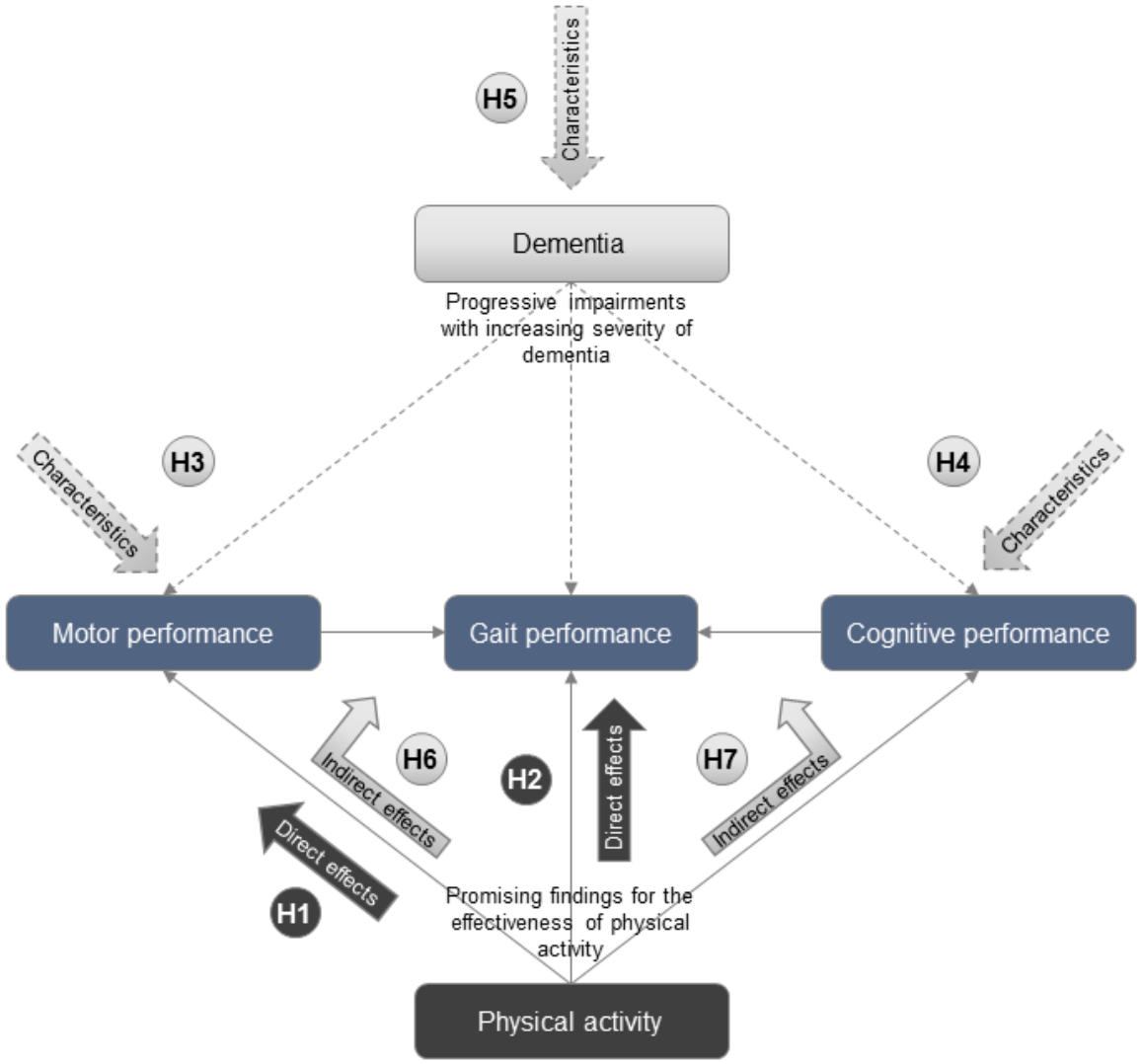


Figure 3. Associations and expected effects of physical activity on motor and gait performance in individuals with dementia, considering characteristics of participants and impacts of changes in underlying motor and cognitive performance on changes in gait performance (H: Hypotheses).

The two primary and four secondary research questions, as well as the seven hypotheses, are examined in five research articles presented in Chapters 5 and 6. Manuscripts I and II focus on research questions A1 and A2. By considering current research practices and providing an example for a high-quality study design, manuscript III addresses research question A3, and simultaneously builds an important basis to answer research question B. Manuscript IV considers research question B1 and hypothesis 3. Finally, manuscript V concentrates on research questions B2, B3, and B4 as well as hypotheses 4-7.

5 Establishing a high-quality methodological approach to investigate the effectiveness of physical activity in individuals with dementia

5.1 Qualitative examination of motor assessments for individuals with dementia

Manuscript I

Summary: Inappropriate motor assessments are among the methodological limitations of previous studies (Blankevoort et al., 2010; Hauer et al., 2006; Suttanon et al., 2010). Applied motor assessments usually were developed for cognitively unimpaired older adults and thus are not tailored to specific characteristics and needs of IWD. Accordingly, manuscript I aims to qualitatively analyze motor assessment applied in previous RCT, discuss possible adaptions, and derive recommendations for assessing motor performance in IWD. Herein, it focuses on research question A1 and builds an important basis for establishing a high-quality methodological approach to investigate the effectiveness of physical activity on motor and gait performance in IWD.

Published in: European Review of Aging and Physical Activity

Published on: April 13th, 2019

Reference:

Trautwein, S., Barisch-Fritz, B., Scharpf, A., Bossers, W., Meinzer, M., Steib, S., Stein, T., Bös, K., Stahn, A., Niessner, C., Altmann, S., Wittelsberger, R., & Woll, A. (2019). Recommendations for assessing motor performance in individuals with dementia: Suggestions of an expert panel - a qualitative approach. *European Review of Aging and Physical Activity*, *16*(5). https://doi.org/10.1186/s11556-019-0212-7

Recommendations for assessing motor performance in individuals with dementia: suggestions of an expert panel – a qualitative approach[6]

Sandra Trautwein[1*], Bettina Barisch-Fritz[1], Andrea Scharpf[1], Willem Bossers[2], Marcus Meinzer[3,4], Simon Steib[5], Thorsten Stein[1], Klaus Bös[1], Alexander Stahn[6,7], Claudia Niessner[1], Stefan Altmann[1], Rita Wittelsberger[1] and Alexander Woll[1]

[1] Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

Full list of author information is available at the end of the article

[*] Correspondence: sandra.trautwein@kit.edu

5.1.1 Abstract

Background: Recommendations for assessing motor performance in individuals with dementia (IWD) are rare, and most existing assessment tools previously applied in IWD were initially developed for healthy older adults. However, IWD and their healthy counterparts differ in motor and cognitive capabilities, which needs to be considered when designing studies for this population. This article aims to give recommendations for motor assessments for IWD and to promote standardisation based on a structured discussion of identified assessment tools used in previous trials.

Methods: Appropriateness and standardisation of previously applied motor assessments for IWD were intensively discussed using a qualitative approach during an expert panel. Furthermore, the use of external cues and walking aids, as well as psychometric properties were considered. Starting with a comprehensive overview of current research practice, the discussion was gradually specified and resulted in the elaboration of specific recommendations.

Results: The superior discussion emphasised the need for tailoring motor assessments to specific characteristics of IWD and attaching importance to standardised assessment procedures. Specific recommendations include the use of sequential approaches, which incorporate a gradual increase of complexity from simple to more difficult tasks, a selection of motor assessments showing sufficient relative reliability and appropriateness for IWD, as well as allowing external cues and walking aids when restricted to repeated instructions and commonly used devices, respectively.

[6] Manuscript I is published in a British journal and thus is written in British English. Some minor formal adaptions were made to the version presented in this thesis to ensure uniform formatting.

Conclusions: These are the first recommendations for assessing motor performance in IWD based on a comprehensive qualitative approach. Due to limited evidence, it was not possible to address all existing questions. It is therefore important to evaluate these recommendations in studies with IWD. Besides tailoring and evaluating available assessments, future research should focus on developing specific tools for IWD. Moreover, further progress in standardisation is necessary to enhance comparability between different trials. This article provides initial approaches for overcoming existing limitations in trials with IWD by giving recommendations and identifying future research questions, and therefore contributes to enhancing evidence regarding efficacy and effectiveness of physical activity interventions.

Keywords: Dementia, Physical activity, Exercise, Motor performance, Motor assessments

5.1.2 Background

Designing studies to investigate the efficacy and effectiveness of physical activity on motor and cognitive performance in individuals with dementia (IWD) is challenging. Despite increasing research in this area (Ahlskog et al., 2011), there still is insufficient evidence, which can be explained by methodological limitations, unspecific interventions, or inappropriate assessments (Brett et al., 2016; Gonçalves et al., 2018; Hauer et al., 2006). This emphasises the need for further high quality studies guided by suggestions for optimised interventions and sensitive assessment tools.

Previous trials which aimed at improving motor performance in IWD frequently applied interventions and assessment tools not adapted to the target population (Fox, Henwood, Keogh, & Neville, 2016; Hauer et al., 2006). Considering significant differences in cognitive and motor performance (Allan et al., 2005; Baddeley et al., 1986; Manckoundia et al., 2006; Perry & Hodges, 1999; van Iersel et al., 2004), however, it is not possible to directly translate study designs developed for healthy older adults to IWD. Especially interventions and assessments need to be tailored to the specific characteristics of IWD, such as decreased executive functioning, reduced attentional and memory capacities (Baddeley et al., 1986; Perry & Hodges, 1999), diminished ability to develop and perform new complex motor sequences (Hauer & Oster, 2008), and impaired gait and balance performance (Allan et al., 2005; van Iersel et al., 2004). For motor assessments, it is important to consider these cognitive and motor impairments,

as both aspects hamper a successful participation. Besides research, this is also important for clinical practice considering prognostic, diagnostic, and therapeutic reasons (Mancini & Horak, 2010; Zijlstra & Aminian, 2007).

Due to cognitive impairment, IWD are often not able to concentrate on comprehensive or complex, unknown tasks, and frequently show difficulties in comprehending instructions, developing appropriate motor actions, and remembering these during execution (Hauer & Oster, 2008). Considering this, there are indications that cognitive impairment may influence the outcome of motor assessments and compromise their feasibility (Rockwood, Awalt, Carver, & MacKnight, 2000). Appropriate assessments therefore need to be tailored to cognitive impairments. It has been suggested that assessments should be of short duration, use appropriate instructions (e.g. no verbal focus, demonstration of the task, clear, short, and repeated instructions; Hauer & Oster, 2008), and include simple motor tasks. Moreover, previous studies discuss the use of external cues thereby considering decreased cognitive abilities. For example, van Iersel, Benraad, and Olde Rikkert (2007) emphasise the need for external cues, which are required to ensure feasibility and to achieve high relative reliability. In contrast, Hauer and Oster (2008) argue that external cues could influence the results and possibly reflect speed, reliability, and quality of external cues rather than actual performance of IWD.

Considering motor influences, disease-specific gait and balance impairments are often accompanied by age-related degeneration, such as frailty (Hajek et al., 2016). Together, they frequently compromise the ability of IWD to walk independently, to complete more complex balance tasks, and to cover longer distances (Suttanon et al., 2012; van Iersel et al., 2004), which are necessary to participate in many assessments. Thus, appropriate assessments also need to be tailed to motor impairments. In this context, it is important to discuss the use of walking aids during assessments. IWD are frequently dependent on walking aids (van Iersel et al., 2007), which ensure participants' safety. However, an investigation in geriatric patients without or with mild to moderate cognitive impairment found that walking aids impede the detection of gait or mobility deficits and thus adversely affect identifying changes over time (Schwenk, Schmidt, Pfisterer, Oster, & Hauer, 2011).

Furthermore, the significance of assessments depends on sound psychometric properties, which are necessary to draw meaningful conclusions (Fox et al., 2016). Investigations determining psychometric properties of motor assessment tools in IWD are rare, especially concerning validity (Bossers, van der Woude, Boersma, Scherder, & van Heuvelen, 2012). Considering test-retest reliability, intraclass correlation coefficients ranging between .42 and .99 and high intra-individual variability were found for most previously applied assessment tools, indicating fair to excellent relative reliability but insufficient absolute reliability (Alencar, Dias, Figueiredo, & Dias, 2012; Blankevoort, van Heuvelen, & Scherder, 2013; Bossers, van der Woude, Boersma, Scherder, & van Heuvelen, 2014; Farrell, Rutt, Lusardi, & Williams, 2010; Fox, Henwood, Neville, & Keogh, 2014; McGough et al., 2013; Ries, Echternach, Nof, & Gagnon Blodgett, 2009; Suttanon, Hill, Dodd, & Said, 2011; Tappen, Roach, Buchner, Barry, & Edelstein, 1997; Thomas & Hageman, 1999; van Iersel et al., 2007; Wittwer, Webster, Andrews, & Menz, 2008). Summarising these results, it has been concluded that the considered motor assessments are appropriate for detecting inter-individual differences in cross sectional or controlled intervention studies, whereas the intra-individual variability was too large to be suitable for investigating intra-individual performance changes (Blankevoort et al., 2013; Ries et al., 2009).

Considering all above-mentioned aspects leads us to the question which motor assessments are actually appropriate for IWD. Unfortunately, there is currently a lack of recommendations or guidelines on how to assess motor performance in IWD. In two of the few studies addressing this issue, Bossers et al. (2012) and McGough et al. (2019) performed a systematic and scoping review, respectively, and recommended using the Berg Balance Scale, the Performance Oriented Mobility Assessment, the Timed Up & Go Test (TUG), short distance walk tests (WT), sit-to-stand (STS) tests, isometric strength assessments, and the 6-min WT for IWD in mild to moderate stages of the disease. Both research groups derived their recommendations based on frequency of use and observed outcome effect sizes in previous trials. They also considered investigations of psychometric properties. However, the authors noted large heterogeneity in the assessment tools used across trials and indicated that insufficient information about feasibility, sensitivity to change, and psychometric properties was frequently provided.

These systematic reviews provide first indications for appropriate motor assessment tools for IWD. However, these recommendations predominantly concern quantitative aspects, and subsequently do not sufficiently consider specific characteristics. There are no comprehensive qualitative approaches focusing on tailoring motor assessments to IWD.

Besides tailoring motor assessments to the specific characteristics of IWD, their standardisation is also important. Gonçalves et al. (2018) note that large heterogeneity in outcomes and assessment tools limits evidence of the efficacy and effectiveness of physical activity in IWD. A detailed review of previous trials also shows that different variations of motor assessment tools were used, which hampers comparability. This emphasises the urgent need to standardise existing motor assessments used in IWD.

This article describes the outcome of a consensus meeting of an international expert panel that aimed to derive recommendations for assessing motor performance in IWD. Due to limited information available for appropriate assessments and their standardisation, the specific goals of the panel were:

1. To discuss the selection of appropriate existing motor assessments, and the standardisation of assessment procedures.
2. To develop standards and procedures for using external cues and walking aids during assessments.
3. To examine psychometric properties of recommended assessments.

5.1.3 Methods

5.1.3.1 Organisation and participants of the expert panel

The international expert panel was organised by, and held at the Karlsruhe Institute of Technology (KIT) at the Institute of Sports and Sports Science (Karlsruhe, Germany). It was composed according to two main aspects: interdisciplinary variety and practical relevance. Interdisciplinarity and methodological variety was achieved by inviting researches from motor and cognitive sciences as well as the humanities, social, and natural sciences. Practical relevance was obtained by involving researchers with direct experience in the development and evaluation of motor assessments or with experience in dementia research. The initial group invited to this expert panel comprised 27 international researchers, who were chosen based on existing relations with the KIT

and a literature screening aimed to identify researchers located within geographical proximity who were interested in motor assessments in IWD.

The expert panel consisted of two interrelated one-day meetings, aiming to achieve an iterative structure. At the first meeting in December 2014, the participants were twelve researchers from five institutions in two countries (Germany and Australia) in the disciplines sports science (especially focusing on biomechanics, human movement science, motor control and learning, sports psychology, sports therapy, and training science), movement and sport gerontology, and psychology. The second meeting in February 2015, comprised a group of five researchers from Germany and the Netherlands covering the disciplines sports science (especially focusing on biomechanics, human movement science, sports therapy, and training science), as well as movement and sport gerontology. With one exception, all researchers of the second meeting also participated in the first. More information on participating researchers is given in the declaration section (see Authors' information).

5.1.3.2 Discussion and guiding questions during the expert panel

Prior to the first meeting, the host institution (KIT) elaborated guiding questions based on challenges in assessing motor performance in IWD identified by literature review. These questions included aspects of appropriateness and standardisation of motor assessments used in previous trials with IWD, use of external cues and walking aids during assessments, and psychometric properties of the considered tools (see Table 9). To enable these guiding questions to be discussed, the KIT research group presented an introductory overview of current research practises and participating researchers introduced their own research experiences with IWD at the first meeting. The discussion was stimulated by the research team of the host institution and one participating researcher supported its guidance.

Table 9. Guiding questions for the discussion during the expert panel

	Guiding questions
Appropriateness and standardisation	Are existing motor assessments appropriate for investigating motor performance in IWD?
	Which motor assessments can be recommended for IWD?
	How can these motor assessments be standardised?

Use of external cues and walking aids	Should the use of external cues during motor assessments in IWD be allowed or not? If yes, which kind of external cues should be allowed? How can external cues be standardised? Should the use of walking aids be allowed during motor assessments in IWD or not? If no, what if IWD are not able to perform task without? If yes, which kind of walking aids should be allowed?
Psychometric properties	Are existing motor assessments valid and reliable to investigate motor performance in IWD?

IWD: individuals with dementia

The first meeting focused on the appropriateness of available motor assessments. In a first step, specific characteristics of IWD and derived demands for motor assessments were considered. Subsequently, these demands were applied to motor assessment tools used in previous trials, which were identified in randomised controlled trials initially analysed for a systematic review assessing effects of physical activity on motor and cognitive performance in IWD (Scharpf, Servay, & Woll, 2013), supplemented by studies of Bossers et al. (Bossers et al., 2015, 2016; Bossers, Scherder et al., 2014; see Table 10). Each tool was then rated whether it sufficiently considered specific characteristics of IWD or could be tailored to them. Specific criteria that impacted this evaluation were duration, instructions, and complexity of each assessment tool, as well as physical strains caused in participants. To address standardisation, descriptions and variations of identified motor assessments were considered and possible modifications were discussed. The discussion concerning the use of external cues and motor assessments focused on the two contrary points of view identified in literature: their need to ensure feasibility and safety vs. the influence of external cues and walking aids on the results of motor assessments. Finally, psychometric properties of the identified motor assessments were considered based on available investigations.

Table 10. Selection of motor assessments [a] discussed in the expert panel

Assessments	Outcomes assessed in previous trials	Description
Balance		
Static balance assessments		
One-leg balance test (Vellas et al., 1997)	Balance (Rolland et al., 2007)	Standing on one leg while participants' ability to maintain this stance for 5 seconds is recorded.
Single leg stance	Lower leg strength and balance (Burgener et al., 2008)	Standing on a single leg alternately for 60 seconds with both eyes open and closed while time is recorded.
Frailty and Injuries: Cooperative Studies of Intervention Technique - subtest 4 (Rossiter-Fornoff, Wolf, Wolfson, & Buchner, 1995)	Balance, static balance (Bossers et al., 2015; Bossers, Scherder et al., 2014)	Performing four different stances for 10 seconds while participants' ability to maintain these stances is evaluated: (1) feet together, (2) semi-tandem, (3) tandem, (4) single-leg.
Posturography platforms assessing postural sway	Balance (Toulotte et al., 2003)	Standing quietly on a posturography platform for 60 seconds with eyes open while elliptical area covered by moving centre of gravity is recorded.
Functional Reach Test (Duncan, Weiner, Chandler, & Studenski, 1990)	Balance and stability (Netz et al., 2007)	Standing next to a wall, holding one arm parallel to a metre stick attached to the wall at shoulder height, and reaching forward as far as possible without losing balance or changing foot position, while distance from starting to end position is recorded.
Dynamic balance assessments		
Figure of Eight Test (Johansson & Jarnlo, 2009)	Balance, dynamic balance (Bossers et al., 2015; Bossers, Scherder et al., 2014)	Walking a lap of a standard figure-eight trajectory as quickly and accurately as possible while walking speed and number of oversteps are recorded.
Groningen Meander Walking Test (Bossers, van der Woude et al., 2014)	Balance, dynamic walking ability (Bossers et al., 2015)	Walking over a meandering curved line as quickly and accurately as possible while time and number of oversteps are recorded. Use of a walking aid is allowed.
Balance scales		
Berg Balance Scale (Berg, 1989)	Functional balance, balance impairment (Burgener et al., 2008; Christofoletti et al., 2008; Fajersztajn, Cordeiro, Andreoni, & Garcia, 2008)	14-item functional balance assessment with simple everyday tasks (reaching, bending, transferring, standing and rising) which are graded on a five-point ordinal scale (0 to 4).
Performance Oriented Mobility Assessment (Tinetti, 1986)	Gait and balance (Hauer et al., 2012; Santana-Sosa et al., 2008)	Scale with two parts, assessing balance (sitting balance, rising from a chair and sitting down, standing balance with eyes open then closed and turning balance) and gait (gait initiation, step length, height, symmetry and continuity, as well as path direction and trunk sway).

Mobility and gait		
Get up & go tests		
Get-Up and Go Test (Mathias, Nayak, & Isaacs, 1986)	Not specified (Rolland et al., 2007)	Standing up from a chair, walking 3 metres, turning around, walking back to the chair, and sitting down, while performance is evaluated from 1 to 5 (1=no instability to 5=very abnormal). Use of a walking aid is allowed.
Timed Up & Go Test (Podsiadlo & Richardson, 1991)	Mobility, functional mobility, balance, dynamic balance, locomotion, muscle-nerve coordination, agility (Bossers et al., 2015; Bossers, Scherder et al., 2014; Christofoletti et al., 2008; Fajersztajn et al., 2008; Hauer et al., 2012; Netz et al., 2007; Toulotte et al., 2003)	Standing up from a chair, walking 3 metres, turning around, walking back to the chair, and sitting down, while time is measured. Use of a walking aid is allowed.
8-foot up-and-go test (Rikli & Jones, 2006)	Speed, agility, and balance while moving (Santana-Sosa et al., 2008)	Standing up from a chair, walking 8 feet, turning around, walking back to the chair, and sitting down, while time is measured.
Manual Timed Up & Go Test (Lundin-Olsson, Nyberg, & Gustafson, 1998; Shumway-Cook, Brauer, & Woollacott, 2000)	Mobility (Fajersztajn et al., 2008)	Timed Up & Go Test with additional manual task (carrying a glass of water).
Cognitive Timed Up & Go Test (Shumway-Cook et al., 2000)	Mobility (Fajersztajn et al., 2008)	Timed Up & Go Test with additional cognitive task (counting backwards by threes).
Walk tests / instrumented gait analysis		
6-metre walk test (Guralnik, Seeman, Tinetti, Nevitt, & Berkman, 1994)	Mobility, walking speed (Bossers et al., 2015; Bossers, Scherder et al., 2014; Rolland et al., 2007)	Walking 6 metres with comfortable pace while time is recorded. Use of walking aid is allowed.
10-metre walk test (Guralnik, Seeman et al., 1994)	Walking speed (Toulotte et al., 2003)	Walking 10 metres with comfortable pace while time is recorded. Use of walking aid is allowed.
Instrumented gait analysis (Bessou, Dupui, Montoya, & Pagès, 1988; Kressig & Beauchet, 2006)	Walking speed, stride length, double limb support time (Kemoun et al., 2010)	Walking at a comfortable pace over an electronic walkway while spatiotemporal gait parameters are recorded.

Strength		
Sit-to-stand tests		
Five Times Sit-to-Stand Test (Csuka & McCarty, 1985)	Lower extremity muscle strength and muscle endurance (Hauer et al., 2012; Netz et al., 2007)	Performing five repetitions of the sit-to-stand task without upper extremity assistance while time is recorded.
30-s chair-stand test (Jones, Rikli, & Beam, 1999)	Muscle dynamic strength endurance of legs (Santana-Sosa et al., 2008)	Performing as many repetitions of the sit-to-stand task as possible within 30 seconds with arms folded across chest.
Modified 30-s chair-stand test, use of upper limbs allowed (Blankevoort et al., 2013; Jones et al., 1999)	Lower body strength, leg strength (Bossers et al., 2015; Bossers, Scherder et al., 2014)	Performing as many repetitions of sit-to-stand task as possible within 30 seconds with upper extremity assistance.
Stair-climbing performance (Reuben & Siu, 1990)	Functional performance (Hauer et al., 2012)	Climbing a flight with 13 stairs while time is recorded.
Instrumented assessments		
Maximal isometric strength assessed with dynamometers (Verkerke et al., 2003)	Maximal isometric muscle strength, maximal knee extension strength (Bossers et al., 2015; Bossers, Scherder et al., 2014; Hauer et al., 2012)	Pushing as hard as possible against the dynamometer after adopting a standardised position while maximum strength and integral over time are recorded.
One-repetition maximum in leg press	Maximal dynamic concentric muscle strength in hip and knee extensors (Hauer et al., 2012)	One-repetition maximum as achieved in the leg-press training machine.
Upper limbs strength		
Handgrip dynamometer	Handgrip strength (Hauer et al., 2012)	Putting maximum force on a dynamometer while maximal handgrip strength is recorded.
Arm curl test (Rikli & Jones, 2006)	Muscle dynamic strength endurance of upper body (Santana-Sosa et al., 2008)	Performing as many biceps curls as possible within 30 seconds holding a hand weight of 5 pounds (women) / 8 pounds (men).
Endurance		
Walk tests		
2-min walk test (Cooper, 1968)	Ambulation (Cott et al., 2002)	Walking for 2 minutes while distance is recorded. Use of usual walking aids is allowed.
6-min walk test (Enright, 2003)	Walking performance (Venturelli et al., 2011)	Walking for 6 minutes without use of walking aids while distance is recorded.
Modified 6-min walk test, use of walking aids allowed (Tappen et al., 1997)	Walking endurance, functional mobility (Bossers et al., 2015; Bossers, Scherder et al., 2014; Tappen et al., 2000)	Walking for 6 minutes while distance is recorded. Use of usual walking aids / physical assistance is allowed.
2-min step test (Rikli & Jones, 2006)	Aerobic endurance (Santana-Sosa et al., 2008)	Performing as many full steps as possible within 2 minutes, raising knees to a point midway between the patella and iliac crest.

Flexibility		
Chair sit-and-reach test (Rikli & Jones, 2006)	Flexibility, flexibility of lower body (Santana-Sosa et al., 2008; Toulotte et al., 2003)	Stretching one leg keeping heel on the floor and trying to touch the toes with the fingers while sitting on a chair while distance between the fingers and toes is recorded.
Back scratch test (Rikli & Jones, 2006)	Flexibility of upper body (Santana-Sosa et al., 2008)	Reaching over the shoulder with one hand and up the middle of the back with the other hand while the distance between extended middle fingers is recorded.
Functional performance		
(Modified) Short Physical Performance Battery (Guralnik, Simonsick et al., 1994)	Functional performance (Hauer et al., 2012)	Assessment battery with three subtests including standing balance (tandem, semi-tandem, and side-by-side stands), walking speed over an 8-foot walking course, and Five Times Sit-to-Stand Test are graded on a five-point ordinal scale (0 to 4). The modified version comprises two subtests including the Five Times Sit-to-Stand Test and gait performance (maximum walking speed, step frequency, cadence).
Senior Fitness Test (Rikli & Jones, 2006)	Functional capacity (Santana-Sosa et al., 2008)	Assessment battery including (1) 30-s chair stand and arm curl test, (2) chair sit-and-reach and back scratch test, (3) 8-foot up-and-go test, and (4) 2-minute step test
Physical Performance Test (Reuben & Siu, 1990)	Performance based motor function activities of daily living (Bossers et al., 2016)	Assessment battery with seven items (writing a sentence, simulated eating, lifting a book onto a shelf, putting on a jacket, picking up a coin from the floor, walking 50 feet, and turning 360°), which are scored on a 4-point Likert scale.
Erlangen Test of Activities of Daily Living (Graessel et al., 2009)	Performance based activities of daily living (Bossers et al., 2016)	Assessment battery with five items (pouring a drink, spreading and cutting a sandwich, opening a small cupboard with a key, washing hands, and tying a bow on a present), which are rated according to correctly performed substeps (0-6 points).

[a] Motor assessments displayed in Table 10 were identified in trials initially analysed for a review assessing effects of physical activity on motor and cognitive performance in IWD (Scharpf et al., 2013), supplemented by studies of Bossers et al. (Bossers et al., 2015, 2016; Bossers, Scherder et al., 2014).

The discussion of all guiding questions was gradually specified from open brain storming to personal estimations and final feedback rounds. Thereby, advantages and disadvantages of individual assessment tools and general procedures were gathered and discussed. In doing so, the first meeting elaborated a comprehensive decision basis for the second meeting and developed specific questions for each assessment tool.

Based on the results of the first meeting, the second meeting aimed to derive specific recommendations in a smaller group setting. It started with a summary of the results,

which were then examined based on common research practice and own research experiences. In the next step, guiding questions were again critically reflected and specific questions for each assessment tool were discussed in detail. As described above, the discussion again was gradually specified. Finally, the appropriateness of assessment tools and use of external cues and walking aids was established by voting. Consensus was defined as an 80 % majority. If consensus was not directly reached, another discussion round was started.

5.1.4 Results

The discussion on current research practices and experiences resulted in a consensus for applying a sequential approach to assess motor performance in IWD. This means that the level of difficulty is gradually increased, starting with simple and proceeding to more complex tasks, if possible. Such an approach takes specific characteristics and needs of the target population into account, and also considers their heterogeneity and reduced physical capacity. This guarantees appropriate requirements tailored to individual performance and improves the feasibility of assessments and comparability of results. For the further discussions, this sequential approach was supposed to be a basic assumption.

5.1.4.1 Recommendations for appropriateness and standardisation of motor assessments

Estimating the appropriateness of motor assessments was performed separately for different physical domains: balance, mobility and gait, strength, endurance, flexibility, and functional performance. To ensure a clear understanding of assessments, short descriptions are given in Table 10.

Based on several underlying considerations extensively described in the following sections, the discussion resulted in recommending the motor assessments summarised in Table 11. As the selection of outcomes and assessments depends on specific objectives and the framework of investigations or aims of clinical examinations, we do not recommend a fixed assessment battery but rather propose a possible selection. Hence, several alternatives are given in Table 11. When composing an assessment battery, it is of great importance to consider the limited capacity of IWD. Thus, we advise restricting the maximum duration of assessments, including rests, to 60 min.

Table 11. Recommended motor assessments

Outcome	Assessments
Balance	Frailty and Injuries: Cooperative Studies of Intervention Technique - subtest 4 (Rossiter-Fornoff et al., 1995) Groningen Meander Walking Test (Bossers, van der Woude et al., 2014) If the investigation/clinical examination focuses on balance: Berg Balance Scale (Berg, 1989) or Performance Oriented Mobility Assessment (Tinetti, 1986)
Mobility and gait	Timed Up & Go Test (Podsiadlo & Richardson, 1991) 6-metre walk test (Guralnik, Seeman et al., 1994) Instrumented gait analysis (GAITRite, comfortable pace, single and dual tasks; Kressig & Beauchet, 2006)
Lower limb strength	Modified 30-s chair-stand test (Blankevoort et al., 2013; Jones et al., 1999), including time for five repetitions
Endurance	With constraints, if endurance is an important outcome: 2-min walk test (Cooper, 1968) or 6-min walk test (Enright, 2003)
Functional performance	Short Physical Performance Battery (Guralnik, Simonsick et al., 1994) Physical Performance Test (Reuben & Siu, 1990)

5.1.4.1.1 Balance assessments

To investigate balance, previous trials applied static (one-leg balance test [Vellas et al., 1997], single leg stance, Frailty and Injuries: Cooperative Studies of Intervention Technique - subtest 4 [FICSIT-4; Rossiter-Fornoff et al., 1995], posturography platforms, and Functional Reach Test [Duncan et al., 1990]) and dynamic approaches (Figure of Eight Test [Johansson & Jarnlo, 2009] and Groningen Meander Walking Test [GMWT; Bossers, van der Woude et al., 2014]). They also utilised balance scales (Berg Balance Scale [Berg, 1989] and Performance Oriented Mobility Assessment [Tinetti, 1986]) encompassing both static and dynamic tasks.

To assess static balance, we recommend the FICSIT-4. Compared to one-leg balance tests, this assessment fulfils the requirement of a sequential approach, starting with a less demanding postural position (parallel stance) which is gradually increased (semi-tandem, tandem, and one-leg stance). Assessment tools recording postural sway, like posturography platforms or force plates, can provide more precise information on static balance. However, many of these assessments cannot be recommended for IWD, since they overtax their physical and balance capabilities. In this context, Ruhe, Fejer, and Walker (2010) performed a systematic review considering various ages and health groups and suggested that three to five repetitions of 90 s each are necessary to reach acceptable reliability values for centre of pressure sway measures. Based on our own

experiences, we thus queried feasibility for the majority of IWD. The same applies for the Functional Reach Test. Assuming a non-satisfactory execution, like not leaning forward as far as possible due to fear of falling (Allan, Ballard, Rowan, & Kenny, 2009), we concluded that the Functional Reach Test is not a valid assessment for static balance in IWD.

The Figure of Eight Test and the GMWT both assess dynamic balance by determining speed and accuracy while walking a prescribed course – a figure of eight trajectory and a meandering curved line, respectively. Compared to the Figure of Eight Test, the walking course of the GMWT was simplified and thereby tailored to the specific characteristics of IWD (Bossers, van der Woude et al., 2014). Therefore, we recommend using the GMWT to assess dynamic balance in IWD.

Compared to single balance assessments, balance scales like the Berg Balance Scale or the Performance Oriented Mobility Assessment include various items predominately focusing on important tasks for everyday life. With small restrictions concerning the tasks of leaning forward and one-leg stance, all items seem to be feasible for IWD. Thus, we consider both balance scales appropriate to use with IWD. Even though these scales evaluate balance more comprehensively than single assessments, their longer duration needs to be considered. Therefore, we recommend balance scales for investigations or clinical examinations focusing on balance, but suggest using single assessments in trials and clinical examinations investigating various physical domains.

5.1.4.1.2 Mobility and gait assessments

Common mobility and gait assessments in IWD include get up and go tests (Mathias et al., 1986), WT (Guralnik, Seeman et al., 1994), and gait analyses (Kressig & Beauchet, 2006). All of these assessments are of short duration, apply simple instructions, and include familiar tasks from everyday life. Based on these estimations, we recommend using all three types of mobility and gait assessments for IWD. However, different variations of get up and go tests and WT exist why standardisation of assessment procedures is very important.

Regarding get up and go tests, different versions vary concerning scoring methods and walking distances. The TUG, a get up and go test version introduced by Podsiadlo and Richardson (1991), allows a quantitative evaluation through timing and is also the most frequently used approach. We therefore recommend using this version of get up and

go tests. However, the TUG and other available get up and go test versions consist of various short tasks, which need to be remembered during execution. Thus, the appropriateness of the TUG for IWD is somewhat limited and predominantly applies for IWD in mild stages of the disease or tailored approaches allowing the use of external cues (see below).

There are WT for different walking distances and paces. Considering spatial limitations and relevance for short distance walking situations in everyday life, we recommend assessing walking at a comfortable pace over a course of six metres. Instrumented gait analysis systems, such as GAITRite® (CIR Systems Inc., Franklin, NJ), can be valuable in providing further detailed information on different spatiotemporal gait parameters. Additionally, dual task conditions can reveal interactions between cognition and gait more clearly (Montero-Odasso et al., 2012; Muir et al., 2012). Besides GAITRite®, which is widely used and has been successfully applied in IWD (Sterke, van Beeck, Looman, Kressig, & van der Cammen, Tischa J M, 2012; Verghese, Wang, Lipton, Holtzer, & Xue, 2007), other instrumented gait analysis systems also might be appropriate, but rarely have been investigated in IWD.

5.1.4.1.3 Strength assessments

Concerning strength outcomes, available tools can be classified as lower limb (STS tests [Csuka & McCarty, 1985], stair-climbing performance [Reuben & Siu, 1990], and instrumented strength assessments) or upper limb strength assessments (handgrip dynamometers and arm curl test [Rikli & Jones, 2006]).

For lower limb strength, we recommend STS tests, in particular a modified 30-s chair-stand test, which allows the use of armrests (Blankevoort et al., 2013; Jones et al., 1999). Although STS performance only partly depends on lower limb strength (Lord, Murray, Chapman, Munro, & Tiedemann, 2002; McCarthy, Horvat, Holtsberg, & Wisenbaker, 2004) and using armrests reduces knee and hip moments (Janssen, Bussmann, & Stam, 2002), it is a functional task, which is of high relevance for everyday life. Moreover, many IWD show reduced physical capacity, and thus may not be able to perform the task without the use of armrests. Compared to the Five Times Sit-to-Stand Test, which records the time required to perform five repetitions, the modified 30-s chair-stand test counts the number of repetitions within 30 s and fulfils the criteria of a sequential approach by allowing each participant to be rated independently of the

number of STS repetitions. Additionally, the time required for five repetitions can be simultaneously assessed for all participants reaching this threshold. However, we do not recommend other lower limb strength assessments without constraints. Although stair-climbing performance is a clinically relevant measure of leg power (Bean, Kiely, LaRose, Alian, & Frontera, 2007), its feasibility may be compromised by practical (availability of standardised flight of stairs) and safety (risk of falling) reasons. With regard to instrumented strength assessments (e.g. dynamometers, isokinetic tools, fitness machines, or other apparatus assessing weights), it has been suggested that such assessments are generally too complex and impractical for assessing large groups (Netz, Ayalon, Dunsky, & Alexander, 2004). Moreover, their suitability for IWD is questionable, as task-specific strength assessments are partly based on complex motion sequences, which are not related to everyday motor experiences, and therefore conflict with the decreased ability to develop and perform new or complex motor sequences of IWD (Hauer & Oster, 2008).

We cannot recommend any of the available assessments for upper limb strength without constraints. Dynamometers assessing handgrip strength were scarcely applied in IWD and first need to prove feasibility. The arm curl test was deemed unsuitable for IWD, because it involves a motion sequence unrelated to everyday life (see instrumented strength assessments).

5.1.4.1.4 Endurance assessments

The 6- or 2-min WT (Cooper, 1968; Enright, 2003), as well as the 2-min step test from the Senior Fitness Test (SFT) (Rikli & Jones, 2006) are available for endurance assessments. We do not recommend these assessments without constraints. All assessments require participants to stand or walk for two or six minutes, respectively. In contrast, IWD often suffer from multiple motor impairments, frequently affecting the performance of standing or walking. Thus, available endurance assessments seem unsuitable for IWD. Consequently, we suggest limiting the use of such assessments only if specifically indicated by the study design or aim of clinical examination. Nonetheless, developing novel, feasible endurance assessments for frail IWD and examining feasibility of existing assessments well-established in other populations, such as ergometer tests, are indicated.

5.1.4.1.5 Flexibility assessments

Only few previous investigations have incorporated flexibility assessments in IWD. Accordingly, the discussion did not consider flexibility assessments in detail. Examples are the chair sit-and-reach test, as well as the back scratch test from the SFT (Rikli & Jones, 2006). As information on their feasibility in IWD is scarcely available, we suggest not using these flexibility assessments, unless flexibility is a central outcome measure.

5.1.4.1.6 Functional performance assessments

Looking at functional performance assessments, previous trials applied the Short Physical Performance Battery (Guralnik, Simonsick et al., 1994), the Physical Performance Test (Reuben & Siu, 1990), the Erlangen Test of Activities of Daily Living (Graessel et al., 2009), and the SFT (Rikli & Jones, 2006). Among these, Freiberger et al. (2012) recommend both the Short Physical Performance Battery and the Physical Performance Test for unimpaired older adults. Both assessments apply tasks relevant to everyday life (e.g., simulated eating, putting on a jacket, standing up from a chair, and walking), and previous trials have demonstrated feasibility in IWD. Thus, our recommendation is to include both assessments to assess functional performance in IWD.

The Erlangen Test of Activities of Daily Living, specifically developed for IWD, is easy and short to administer and includes tasks demonstrating relevance for everyday life (e.g. pouring a drink or washing hands; Graessel et al., 2009). This indicates its appropriateness for IWD. However, it seems to be too easy for individuals with mild dementia (Luttenberger, Schmiedeberg, & Grassel, 2012), and therefore we do not recommend it without constraints. Moreover, we do not recommend the SFT for IWD, because it comprises tasks such as the arm curl, chair sit-and-reach, back scratch, and 2-min step tests that were deemed unsuitable for IWD (for details please see above).

5.1.4.2 Recommendations for the use of external cues and walking aids during assessments

Previous trials assessing motor performance in IWD frequently allowed the use of external cues. However, their influence on results has not yet been well-established. Nevertheless, external cues seem to be important to ensure the feasibility of motor assessments in IWD, and are especially necessary for assessments consisting of many short

tasks, such as the TUG (van Iersel et al., 2007). In this context, we note the heterogeneity in external cues applied across previous trials and emphasise the need for standardisation for comparability reasons. Thus, we suggest allowing the exact repetition of instructions but no other external cues, if not otherwise indicated in the assessment protocol. Moreover, we advise a careful documentation and reporting of used external cues. This recommendation of allowing a restricted use of external cues contributes to tailoring motor assessments to specific characteristics of IWD, and is a first step towards standardisation, which needs to be further substantiated. However, the use of external cues is not appropriate for assessments determining complex motor-cognitive performance, such as activities of daily living.

Walking aids are frequently required by both older adults and IWD, and assessment protocols do often not indicate how to deal with them (Schwenk et al., 2011). Despite their influence on detecting gait changes over time (Schwenk et al., 2011), we recommend allowing the use of walking aids applied in everyday life due to safety reasons and to avoid missing data. This may also increase ecological validity, since IWD who use a walking aid in everyday life would be examined in their daily situation. Whenever possible, however, the TUG should be performed without walking aids. Focusing on standardisation, we further suggest restricting the use of walking aids to commonly used aids (e.g. walkers, canes, and crutches), which does not include personal assistance. We also recommend carefully documenting and reporting the use of waking aids. Addressing comparability between baseline and post assessment values, additional qualitative analyses may be indicated when considerable changes in the use of walking aids occurred.

5.1.4.3 Psychometric properties

Only few investigations examining psychometric properties of motor assessments in IWD were available at the time of the two meetings. Thus, there was no profound empirical basis for evaluating psychometric properties of motor assessments, which emphasises the urgent need for further investigations.

Considering available investigations, validity has been examined too scarcely and heterogeneously to draw comprehensive conclusions. As indicated in Table 12, intraclass correlation coefficient values for recommended assessment tools ranged between .57 and .99, reflecting sufficient relative reliability, whereas higher intra-individual variability

shows lower absolute reliability. Based on these findings, it was concluded that the above-recommended assessments show sufficient reliability to assess inter-individual differences in cross sectional or controlled intervention studies, but are not suitable for determining intra-individual changes (Blankevoort et al., 2013; Bossers, van der Woude et al., 2014; Farrell et al., 2010; Fox et al., 2014; McGough et al., 2013; Ries et al., 2009; Suttanon et al., 2011; Tappen et al., 1997; Thomas & Hageman, 1999; van Iersel et al., 2007; Wittwer et al., 2008).

Table 12. Relative (ICC) and absolute (SEM, MDC_{95}) reliabilities for recommended assessments

Recommended assessments		**Test-retest reliability** [a]
Frailty and Injuries: Cooperative Studies of Intervention Technique - subtest 4 (Rossiter-Fornoff et al., 1995)		ICC=.79-.82 (Blankevoort et al., 2013) SEM=.55-.60 points (Blankevoort et al., 2013) MDC_{95}=1.52-1.66 points (Blankevoort et al., 2013)
Groningen Meander Walking Test (Bossers, van der Woude et al., 2014)	Time	ICC=.93-.96 (Bossers, van der Woude et al., 2014) SEM=1.93 s (Bossers, van der Woude et al., 2014) MDC_{95}=5.35 s (Bossers, van der Woude et al., 2014)
	Oversteps	ICC=.57-.79 (Bossers, van der Woude et al., 2014) SEM=1.58 oversteps (Bossers, van der Woude et al., 2014) MDC_{95}=4.38 oversteps (Bossers, van der Woude et al., 2014)
Berg Balance Scale (Berg, 1989)		N/A
Performance Oriented Mobility Assessment (Tinetti, 1986)		ICC=.96 (van Iersel et al., 2007)
Timed Up & Go Test (Podsiadlo & Richardson, 1991)		ICC=.76-.99 (Blankevoort et al., 2013; Ries et al., 2009; Suttanon et al., 2011; Thomas & Hageman, 1999)[b] SEM=1.43-3.03 s (Blankevoort et al., 2013; Ries et al., 2009; Suttanon et al., 2011)[b] MDC_{95}=2.42-8.07 s (Blankevoort et al., 2013; Ries et al., 2009; Suttanon et al., 2011)[b]
6-metre walk test (Guralnik, Seeman et al., 1994)	Time	ICC=.92 (Thomas & Hageman, 1999)
	Speed	ICC=.83-.89 (Blankevoort et al., 2013) SEM=.09-.11 m/s (Blankevoort et al., 2013) MDC_{95}=.25-.29 m/s (Blankevoort et al., 2013)
	Steps	ICC=.80 (Thomas & Hageman, 1999)
Instrumented gait analysis (GAITRite®, comfortable pace, single task; Kressig & Beauchet, 2006)	Speed	ICC=.95-.98 (McGough et al., 2013; Ries et al., 2009; Wittwer et al., 2008) SEM=0.06 m/s (Ries et al., 2009) MDC_{95}=.11-.13 m/s (Ries et al., 2009; Wittwer et al., 2008)

	Step/ stride length	ICC=.97-.98 (McGough et al., 2013; Wittwer et al., 2008) MDC_{95}=4.15-5.27 / 8.12-10.24 cm (Wittwer et al., 2008)
	Step width	ICC=.92-.95 (Wittwer et al., 2008) MDC_{95}=1.83-2.23 cm (Wittwer et al., 2008)
	Stance/ swing time	ICC=.87-.96 (McGough et al., 2013; Wittwer et al., 2008) MDC_{95}=.03-.06 s (Wittwer et al., 2008)
	Cadence	ICC=.88-.91 (McGough et al., 2013; Wittwer et al., 2008) MDC_{95}=7.64-8.13 steps/min (Wittwer et al., 2008)
Instrumented gait analysis (GAITRite®, comfortable pace, dual task; Kressig & Beauchet, 2006)		N/A
Modified 30-s chair-stand test (Blankevoort et al., 2013; Jones et al., 1999)		ICC=.79-.88 (Blankevoort et al., 2013) SEM=.83-1.52 repetitions (Blankevoort et al., 2013) MDC_{95}=2.30-4.21 repetitions (Blankevoort et al., 2013)
Short Physical Performance Battery (Guralnik, Simonsick et al., 1994)		ICC=.88 (Fox et al., 2014)
Physical Performance Test (Reuben & Siu, 1990)		ICC=.90 (Farrell et al., 2010)
2-min walk test (Cooper, 1968)		N/A
6-min walk test (Enright, 2003)		ICC=.75-.99 (Ries et al., 2009; Tappen et al., 1997)[b] SEM=19.57-21.86 m (Ries et al., 2009)[c] MDC_{95}=39.76 m (Ries et al., 2009)[c]

[a] Between-day test-retest reliability, if not otherwise indicated;

[b] Between-day and within-day test-retest reliability;

[c] Within-day test-retest reliability

ICC: intraclass correlation coefficient; SEM: standard error of measurement; MDC_{95}: minimal detectable change with 95 % confidence interval

5.1.5 Discussion

The present article conveys a consensus of recommendations for assessing motor performance in IWD, which was reached during an expert panel with two interrelated one-day meetings at the KIT in December 2014 and February 2015. These recommendations focus on the appropriateness and standardisation of motor assessments for IWD, deal with the use of external cues or walking aids, and consider psychometric properties of recommended assessments.

To appropriately address IWD, we recommend using a sequential approach and suggest a selection of eight motor assessments to investigate balance (FICSIT-4 and GMWT), mobility and gait (TUG, 6-m WT, and instrumented gait analysis), lower limb strength (modified 30-s chair-stand test), and functional performance (Short Physical Performance Battery and Physical Performance Test). Moreover, we put emphasis on a standardised assessment procedure to ensure comparability between different trials/clinical examinations and to thereby allow conclusions to be drawn based on sound evidence. Considering standardisation in general, we advise allowing a restricted use of external cues and walking aids, and to carefully document and report their use. Psychometric properties could not be considered in-depth, but available investigations determined sufficient relative reliability for the majority of recommended assessments. These recommendations were primarily elaborated for research but equally can be applied in clinical practice. However, lower absolute reliability needs to be considered when assessing intra-individual changes.

To our knowledge, this is the first article giving comprehensive recommendations for assessing motor performance in IWD using a qualitative approach. The few available investigations also focusing on recommendations of motor assessments in IWD analysed assessments used in previous trials from a quantitative perspective, and did not deal with standardisation of assessment procedures or tailoring assessments to specific characteristics of IWD. For the most part, the assessments recommended in this paper coincide with these recommendations (see Bossers et al., 2012 and McGough et al., 2019).

A major strength of the expert panel was the comprehensive and thorough analysis of the appropriateness of motor assessments considering specific characteristics of IWD. Following the expert panel, the recommended assessments were applied in a trial of our own with IWD in mild to moderate stages of the disease (Trautwein, Scharpf, Barisch-Fritz, Niermann, & Woll, 2017) and demonstrated feasibility. This is in line with previous investigations successfully utilising these assessments or determining their reliability. Nevertheless, information on psychometric properties in many cases is still insufficient and further research is needed (Bossers et al., 2012). Furthermore, our own experiences showed that it was not possible to identically adopt assessment procedures common for healthy older adults for IWD. For example, it was necessary to allow external cues in form of repeated instructions. This clearly illustrates the need for

tailored versions of existing motor assessments, which first need to be standardised and evaluated. Unfortunately, it was not possible to discuss more recent findings within another expert meeting.

Potential biases need to be stated concerning the choice of motor assessments. Despite applying a systematic approach, considered assessments were restricted to those utilised in randomised controlled trials with IWD. Thus, other potentially appropriate assessments may be missing. Additionally, the derived recommendations could be biased by the researchers' experiences and preferences. Moreover, this article only considers existing assessments used in previous trials, whereby the recommendations only include the most suitable of the available possibilities. Besides investigating psychometric properties of existing assessments and developing tailored standardised versions, which consider specific characteristics of IWD, future research should also focus on developing new assessments specifically for IWD. In summary, the recommendations in this article were thoroughly deduced from existing literature and consider the psychometric properties as much as possible. However, they should be used carefully as it is important to first evaluate them in different studies with IWD and address further questions due to limited evidence.

5.1.6 Conclusions

This article contributes to giving recommendations on performing motor assessments in IWD. However, these recommendations show a preliminary character and are not able to deal with all existing questions. One main problem is that most assessments applied in previous trials were not developed initially for IWD and are not well-investigated within this target population.

Finally, we indicate the need for further studies investigating common motor assessments for administration in IWD. We further encourage tailoring assessment procedures and evaluating existing motor assessments according to the special characteristics of IWD, and then investigating these adapted versions. Nevertheless, it will be important to develop and investigate specific assessments specifically for IWD, such as the GMWT.

In line with Gonçalves et al. (2018), we encourage scientists and clinical practitioners to reach a consensus concerning the use of motor assessments, and to apply a standardised assessment procedure aiming to enhance comparability in the research field

and clinical practice. With regard to scientific publications, we therefore ask scientists to give a detailed report on how they perform motor assessments in IWD, as different modifications exist and it is often not clear which has been applied.

All efforts undertaken to develop and apply standardised and reliable motor assessments which are appropriate and meaningful for IWD are important steps to enhance evidence concerning efficacy and effectiveness of physical activity on motor performance in IWD.

Abbreviations

FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Technique -subtest 4; GMWT: Groningen Meander Walking Test; IWD: individuals with dementia; KIT: Karlsruhe Institute of Technology; SFT: Senior Fitness Test; STS: sit-to-stand; TUG: Timed Up & Go Test; WT: walk test, walk tests

Declarations

Acknowledgements

We would like to thank Prof. Dr. Dr. h.c. em. Klaus Willimczik for participating in the expert panel and supporting the guidance of the discussion, Dr. Tim Fleiner for participating as well as presenting an experience report regarding daily and research work with IWD, and Lisa Peterson for her linguistic assistance on behalf of the authors.

We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Karlsruhe Institute of Technology.

Funding

This project is financially supported by the Dietmar Hopp Stiftung (St. Leon-Rot, Germany). The sponsor does not have any role in the design of the study, neither in its execution, the collection, analysis, or interpretation of data, the decision to submit results, nor in writing the report.

Availability of data and materials

Not applicable.

Authors' contributions

ST, ASC and AW prepared the expert panel. ST and SS extensively discussed existing motor assessments and their performance in preparation of the expert panel. ST, ASC, WB, MM, TS, KB, AST, CN, SA, RW, and AW participated in the expert panel. ST, ASC, MM, and WB gave presentations within the expert panel. All authors contributed to the development of the assessment battery and further recommendations. ST and BB wrote the manuscript. All authors provided critical feedback and approved the final manuscript.

Authors' information

ST, BB, ASC, WB, MM, and AW performed comprehensive research with IWD and are familiar in working with the target group. ST and WB intensively analysed literature towards research on motor assessments in IWD. WB, SS, TS, KB, CN, SA, and AW acquired expertise in developing and evaluating motor assessments for various populations. ST, BB, ASC, WB, MM, SS, KB, AST, RW, and AW were involved in ageing research. ST, BB, ASC, WB, MM, TS, AST, and AW studied the effects of exercise on cognition and brain plasticity.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1] Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Engler-Bunte-Ring 15, 76131 Karlsruhe, Germany. [2] Center for Human Movement Sciences, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [3] Centre for Clinical Research, University of Queensland, Brisbane, Australia. [4] Department of Neurology, Greifswald University Medicine, Greifswald, Germany. [5] Department of Sport Science and Sport, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany. [6] Unit of Experimental Psychiatry, Department of Psychiatry, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, USA. [7] Institute of Physiology, Charité – Universitätsmedizin Berlin, a corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany.

Publisher's Note

Received: 19 December 2018, Accepted: 31 March 2019, Published online: 13 April 2019

References

All references of manuscript I are included in the List of References at the end of this thesis.

5.2 Quantitative examination of motor assessments for individuals with dementia

Manuscript II

Summary: With respect to quantitative aspects, valid, reliable, and sensitive motor assessments are required to investigate the effectiveness of physical activity on motor and gait performance in IWD (Blankevoort et al., 2010; Hauer et al., 2006). However, the psychometric properties of motor assessments applied in previous RCT are not thoroughly examined in IWD (Blankevoort et al., 2010). Thus, manuscript II aims to summarize and analyze findings on psychometric properties, frequency of use, and effect sizes of motor assessment applied in previous RCT, while considering severity and etiology of dementia, as well as the use of external cues. Based on knowledge gained it established recommendations on motor assessments for IWD. By quantitatively analyzing motor assessments for IWD manuscript II addresses research question A2.

Published in: European Review of Aging and Physical Activity

Published on: November 3rd, 2019

Reference:

Trautwein, S., Maurus, P., Barisch-Fritz, B., Hadzic, A., & Woll, A. (2019). Recommended motor assessments based on psychometric properties in individuals with dementia: A systematic review. *European Review of Aging and Physical Activity*, *16*(20). https://doi.org/10.1186/s11556-019-0228-z

Recommended Motor Assessments Based on Psychometric Properties in Individuals with Dementia: A Systematic Review[7]

Sandra Trautwein[1*], Philipp Maurus[1,2], Bettina Barisch-Fritz[1], Anela Hadzic[1], Alexander Woll[1]

[1] Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

Full list of author information is available at the end of the article

[*] Correspondence: sandra.trautwein@kit.edu

5.2.1 Abstract

Background: Motor assessments are important to determine effectiveness of physical activity in individuals with dementia (IWD). However, inappropriate and non-standardised assessments without sound psychometric properties have been used. This systematic review aims to examine psychometric properties of motor assessments in IWD combined with frequency of use and effect sizes and to provide recommendations based on observed findings.

We performed a two-stage systematic literature search using Pubmed, Web of Science, Cochrane Library, ALOIS, and Scopus (inception - July/September 2018, English and German). The first search purposed to identify motor assessments used in randomised controlled trials assessing effectiveness of physical activity in IWD and to display their frequency of use and effect sizes. The second search focused on psychometric properties considering influence of severity and aetiology of dementia and cueing on test-retest reliability. Two reviewers independently extracted and analysed findings of eligible studies in a narrative synthesis.

Results: Literature searches identified 46 randomised controlled trials and 21 psychometric property studies. While insufficient information was available for validity, we observed sufficient inter-rater and relative test-retest reliability but unacceptable absolute test-retest reliability for most assessments. Combining these findings with frequency of use and effect sizes, we recommend Functional Reach Test, Groningen Meander Walking Test (time), Berg Balance Scale, Performance Oriented Mobility Assessment, Timed Up & Go Test, instrumented gait analysis (spatiotemporal parameters), Sit-to-Stand assessments (repetitions>1), and 6-minute walk test. It is important to consider

[7] Manuscript II is published in a British journal and thus is written in British English. Some minor formal adaptions were made to the version presented in this thesis to ensure uniform formatting.

that severity and aetiology of dementia and cueing influenced test-retest reliability of some assessments.

Conclusion: This review establishes an important foundation for future investigations. Sufficient relative reliability supports the conclusiveness of recommended assessments at group level, while unacceptable absolute reliability advices caution in assessing intra-individual changes. Moreover, influences on test-retest reliability suggest tailoring assessments and instructions to IWD and applying cueing only where it is inevitable. Considering heterogeneity of included studies and insufficient examination in various areas, these recommendations are not comprehensive. Further research, especially on validity and influences on test-retest reliability, as well as standardisation and development of tailored assessments for IWD is crucial.

This systematic review was registered in PROSPERO (CRD42018105399).

Keywords: Physical performance measurements, Cognitive impairment, Validity, Reliability, Frequency of use

5.2.2 Background

Physical activity has gained importance as therapeutic strategy for individuals with dementia (IWD), and in accordance, the number of trials investigating its effectiveness on motor and cognitive performance in IWD has increased (Ahlskog et al., 2011). However, methodological limitations, such as inappropriate or inconclusive motor assessments, affect the derivation of evidence. Thus, further high quality investigations are required (Brett et al., 2016; Gonçalves et al., 2018; Hauer et al., 2006).

Considering motor assessments, high quality is reflected by appropriateness for the intended population, sensitivity to change, sound psychometric properties, and standardisation (Gonçalves et al., 2018; Ries et al., 2009; Wittwer, Webster, & Hill, 2013). In many cases, motor assessments used in previous trials failed to meet these criteria. The majority of applied assessments has predominately been developed for healthy older adults and does not consider specific characteristics of IWD (Fox et al., 2016). However, IWD and unimpaired individuals differ in their cognitive and motor performance (Allan et al., 2005; Baddeley et al., 1986; Manckoundia et al., 2006; Perry & Hodges, 1999; van Iersel et al., 2004). Thus, tailoring motor assessments to IWD is essential to ensure appropriateness. Furthermore, insufficient or inconsistent research

regarding sensitivity to change and psychometric properties in IWD (Bossers et al., 2012) restricts the derivation of meaningful conclusions from applied motor assessments (Muir-Hunter, Graham, & Montero-Odasso, 2015; Telenius, Engedal, & Bergland, 2015b). Referring to this, literature indicates that dementia affects reliability (Blankevoort et al., 2013; Hauer & Oster, 2008; Phillips, Chu, Morris, & Hawes, 1993; Ries et al., 2009), which was scarcely considered in previous trials. With regard to standardisation, previous research utilised a variety of motor assessments and modifications, affecting comparability (Bossers et al., 2012; Gonçalves et al., 2018). Therefore, inappropriateness, insensitivity, inconclusiveness, and non-standardisation limit the derivation of evidence.

Considering heterogeneous cognitive and motor impairments (Allan et al., 2005; Cohen-Mansfield, 2000), motor assessments may not be equally suitable for all IWD. Severity and aetiology of dementia, which are important determinants contributing to this heterogeneity (Cohen-Mansfield, 2000; Valkanova & Ebmeier, 2017), potentially influence psychometric properties of motor assessments. Particularly, test-retest reliability may decrease with increasing severity of dementia, due to growing intra-individual variability or progressive difficulties to participate in motor assessments (Blankevoort et al., 2013; Hauer & Oster, 2008; Phillips et al., 1993; Ries et al., 2009). Similarly, aetiology of dementia can influence test-retest reliability as cognitive and motor impairments vary in time of occurrence and severity in different aetiologies (Cohen-Mansfield, 2000; Muir-Hunter et al., 2015). Moreover, the influence of external cues on test-retest reliability, which are used to compensate for cognitive and motor impairments, has been discussed (Hauer & Oster, 2008; van Iersel et al., 2007).

Literature comprehensively addressing motor assessments for IWD is limited. The importance of research in this area is highlighted in a qualitative approach (Trautwein, Barisch-Fritz et al., 2019) of analysing the appropriateness of motor assessments for IWD. Additionally to elaborating recommendations, this article emphasises the need for tailoring and standardising motor assessments for IWD (Trautwein, Barisch-Fritz et al., 2019). Moreover, three systematic reviews (Bossers et al., 2012; Fox et al., 2016; H.-S. Lee & Park, 2017) and one scoping review (McGough et al., 2019) examined frequency of use, sensitivity to change, and psychometric properties. Bossers et al. (2012) and McGough et al. (2019) identified eight frequently applied, sensitive assessments, showing good to excellent relative test-retest reliability. Fox et al. (2016) found

appropriate relative test-retest reliability, but insufficient absolute test-retest reliability and limited information on validity for several motor assessments. While H.-S. Lee and Park (2017) determined similar intraclass correlation coefficients (ICC), they applied a more stringent rating, suggesting acceptable relative test-retest reliability only for the Berg Balance Scale (BBS). Additionally, they considered the influence of different aetiologies of dementia on relative test-retest reliability, but were not able to draw conclusions due to insufficient research. In summary, these reviews provide an important basis, but do not actually allow a comprehensive quantitative evaluation of motor assessments for IWD. Previous reviews focused on frequency of use and sensitivity to change (Bossers et al., 2012; McGough et al., 2019) or just considered relative reliability and neglected other psychometric properties such as absolute reliability or validity (Bossers et al., 2012; H.-S. Lee & Park, 2017; McGough et al., 2019). They only investigated psychometric properties of the most common motor assessments without taking into account the influences of the heterogeneity of IWD (Bossers et al., 2012; Fox et al., 2016; McGough et al., 2019) or considering further outcomes such as frequency of use or sensitivity to change (Fox et al., 2016; H.-S. Lee & Park, 2017). Moreover, information on how psychometric properties were graded was rare (Bossers et al., 2012; H.-S. Lee & Park, 2017; McGough et al., 2019), no specific recommendations were suggested (Fox et al., 2016; H.-S. Lee & Park, 2017), and the results of different outcomes were not combined when drawing conclusions (Fox et al., 2016). Finally, previous randomised controlled trials (RCT) with IWD applied additional motor assessments which were not considered in previous reviews (Bossers et al., 2012; Fox et al., 2016; H.-S. Lee & Park, 2017; McGough et al., 2019).

With respect to these limitations, we indicated the following main research gaps: (a) comprehensive quantitative approaches combining outcomes of identified reviews including psychometric properties, frequency of use, and effect sizes of motor assessments applied in previous trials with IWD and (b) research on the influence of severity and aetiology of dementia and cueing on test-retest reliability. Therefore, the objectives of this systematic review are: (1) to quantitatively examine motor assessments for IWD used in previous RCT by comprehensively analysing psychometric properties (primary outcome), frequency of use, and effect sizes of those assessments (secondary outcomes) and (2) to assess the influence of severity and aetiology of dementia and cueing on test-retest reliability. Based on primary and secondary outcomes, this review

derives recommendations, which contribute to create consensus and decrease heterogeneity of motor assessments for future research. It needs to be considered that there are several purposes and reasons for applying motor assessments. Motor assessments are essential for diagnostic purposes and to assess changes over time, e.g. in RCT. Regarding specific reasons, they are utilised to determine actual motor performance, but also to evaluate related outcomes, such as frailty (Lundin-Olsson et al., 1998) and risk of falls (McGough et al., 2013), or to draw conclusions on underlying cognitive performance (Beauchet et al., 2008). This review focuses on motor assessments to assess changes over time, but does not further differentiate between various reasons for the use of motor assessments. Instead, it aims to provide a general overview.

5.2.3 Methods

For this systematic review, we considered the guidelines and recommendations of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Statement (Liberati et al., 2009; Moher, Liberati, Tetzlaff, & Altman, 2009). Furthermore, we registered the systematic review in PROSPERO (CRD42018105399).

We performed a two-stage literature search to address the objectives of this systematic review. A first search focused on the identification of motor assessments applied in RCT in IWD. Based on these findings, a second search (main search) aimed to determine publications examining psychometric properties of the identified motor assessments. This approach ensures to focus on those motor assessments commonly applied in IWD and allows the determination of various outcomes required for a comprehensive quantitative evaluation of motor assessments for IWD. The taxonomy of COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative (Mokkink et al., 2010) provided the terminology and definitions of psychometric properties. In line with literature, we applied the terms relative and absolute reliability for reliability and measurement error, respectively (Bruton, Conway, & Holgate, 2000). Relative reliability, quantified by correlation coefficients, refers to the degree to which individual measurements maintain their position within a sample over repeated assessments, while absolute reliability, quantified by standard error of meas-

urements or minimal detectable changes, is the degree to which individual measurements vary over repeated assessments (Bruton et al., 2000; Carter, Lubinsky, & Domholdt, 2013; Ries et al., 2009).

5.2.3.1 First search

For the first search, we examined the electronic databases Pubmed, Web of Science, Cochrane Library, and ALOIS between December 2016 and July 2018 without date restrictions. We applied terms related to dementia, physical activity, and motor performance to identify eligible trials (see Additional file 1 for complete search term), supplemented by manually checking references of indicative articles and reviews. Two reviewers independently screened titles and abstracts (ST and BB) and checked inclusion criteria during full-text analysis (ST and AH). Trials were eligible if they met the following criteria: (a) designed as (cluster) RCT, (b) included individuals with primary dementia (Alzheimer's disease [AD], vascular dementia, frontotemporal dementia, and Lewy body disease) older than 65 years, (c) applied physical activity interventions[8], (d) used motor assessments independent of intended reasons, and (e) were published and written in English or German. We excluded comments, conference abstracts, protocols, and trial registrations. If there were disagreements, the two reviewers consulted a third reviewer (AW) to reach a consensual decision.

One reviewer (ST) extracted the following data from included RCT using a standardised extraction form: sample size, sample characteristics, motor assessments, means and standard deviations of baseline and post motor assessments, corresponding F/t statistics, and effect sizes. A second reviewer (AH) checked the outcomes. The two reviewers discussed ambiguities and disagreements in consensus meetings and consulted a third reviewer (BB) if they reached no agreement.

In addition to analysing frequency of use of identified motor assessments, we calculated time*group interaction effect sizes to represent their sensitivity to change. We determined Cohen's d if F (time*group interaction) or t (between group baseline-post differences) statistics, or baseline-post differences including standard deviations were provided (Thalheimer & Cook, 2002; formulas see Additional file 2). A Cohen's d of

[8] defined as all types of physical activity that are planned, structured, repetitive, and purposive aiming to improve or maintain one or more components of physical fitness (Caspersen, Powell, and Christenson (1985))

0.2, 0.5, and 0.8 represents a small, medium, and large effect size, respectively (Jacob Cohen, 1988). Furthermore, we considered time*group interaction effect sizes provided in RCT.

This first search primarily aimed to identify motor assessments used in previous RCT with IWD and served as basis for the main search. Hence, we did not assess risk of bias.

5.2.3.2 Main search

For the main search, we examined the electronic databases PubMed, Web of Science, Cochrane Library, and Scopus (no date restrictions) between August and September 2018 for terms related to dementia, psychometric properties, and motor assessments identified in the first search (see Additional file 3 for complete search term). Additionally, we manually checked reference lists of indicative articles. Two reviewers (ST and PM) independently screened titles and abstracts and checked inclusion criteria during full-text analysis. Trials were eligible if they fulfilled the following criteria: (a) examined psychometric properties (content validity, construct validity, criterion validity, internal consistency, intra-rater reliability, inter-rater reliability, test-retest reliability, relative and absolute reliability) of (b) motor assessments in (c) individuals with primary dementia (AD, vascular dementia, frontotemporal dementia, and Lewy body disease) aged above 65 years, (d) applied Mini-Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975), and (e) were written and published in English or German. We excluded comments and conference abstracts. The two reviewers discussed disagreements and consulted a third reviewer (BB) to resolve remaining discrepancies.

Two reviewers (ST and PM) independently extracted the following information from eligible investigations utilising a standardised data extraction form: sample size, sample characteristics, motor assessments, methodologies, and statistics of psychometric properties. Moreover, they independently assessed risk of bias of individual investigations with the COSMIN checklist (Mokkink, Vet et al., 2018; Prinsen et al., 2018). The two reviewers resolved disagreements through discussion and consulted a third reviewer (BB) if necessary.

Afterwards, we analysed findings of eligible investigations in a systematic narrative synthesis and summarised extracted information. In order to allow comparability of minimal detectable change values, we calculated percentage minimal detectable

changes at 95 % confidence interval ($MDC_{95\%}$) if any standard error of measurement or minimal detectable change was reported (Portney & Watkins, 2015; Schwenk, Gogulla, Englert, Czempik, & Hauer, 2012; formulas: see Additional file 4).

Moreover, we rated the results of each study against the COSMIN criteria for good measurement properties (Mokkink, Prinsen et al., 2018). Since information on minimal important change of considered motor assessments in IWD is rare (Blankevoort et al., 2013), and no other firm criteria for acceptable values (Smidt et al., 2002) are available, we considered a $MDC_{95\%}$ higher than 30 % as unacceptable (Huang et al., 2011; H.-S. Lee, Park, & Chung, 2017). Based on COSMIN reliability criteria for good measurement properties (Mokkink, Prinsen et al., 2018, p. 28) and indications for unacceptable values (Huang et al., 2011; H.-S. Lee et al., 2017), we rated relative and absolute reliability as follows:

- sufficient relative/absolute reliability (+): ICC≥0.70/minimal detectable change at 95 % confidence interval<minimal important change
- indeterminate relative/absolute reliability (?): ICC not reported/minimal important change not defined
- insufficient relative/absolute reliability (-): ICC<0.70/minimal detectable change at 95 % confidence interval>minimal important change
- unacceptable absolute reliability (↓): $MDC_{95\%}$>30 %

Subsequently, we summarised overall evidence and graded quality of evidence using the Grading of Recommendations Assessment, Development, and Evaluation approach, which considers risk of bias, inconsistency, imprecision, and indirectness of included investigations (Mokkink, Prinsen et al., 2018; Schünemann, Brożek, Guyatt, & Oxman, 2013). Additionally, we analysed the influence of severity and aetiology of dementia and cueing on test-retest reliability. Therefore, we determined severity of dementia according to reported MMSE values (mild: MMSE=26-17, moderate: MMSE=17-10, severe: MMSE<10; Feldman & Woodward, 2005; Forbes et al., 2015; Hogan et al., 2007) and/or classification of publications if range of MMSE was not reported. Due to insufficient information on aetiology, we were only able to compare between AD and various or not reported types. In accordance with Muir-Hunter et al. (2015, p. 257) we defined cueing as “providing any additional verbal, visual, or tactile direction necessary to ensure correct performance of the task after the initial set of

standardized instructions was given". To investigate its influence on test-retest reliability, we classified cueing in five categories, considering information in identified psychometric property studies: (a) not reported, (b) no cueing, (c) verbal cueing, (d) verbal and visual/tactile cueing, and (e) more extensive cueing than (c) and (d) including physical assistance.

5.2.4 Results

5.2.4.1 Systematic searches (first and main search)

The first search revealed 5007 publications. After removing duplicates and initial screening on titles and abstracts, we screened the full texts of 309 publications and included 46 RCT for further analysis. For the main search, we obtained 902 publications. Removing duplicates and initial screening on titles and abstracts yielded 68 publications, of which we scanned full texts. Eventually, we included 21 eligible investigations in the narrative data synthesis (see Figure 4, further information on study characteristics and data extractions are provided in Additional files 5, 6, 7 and 8).

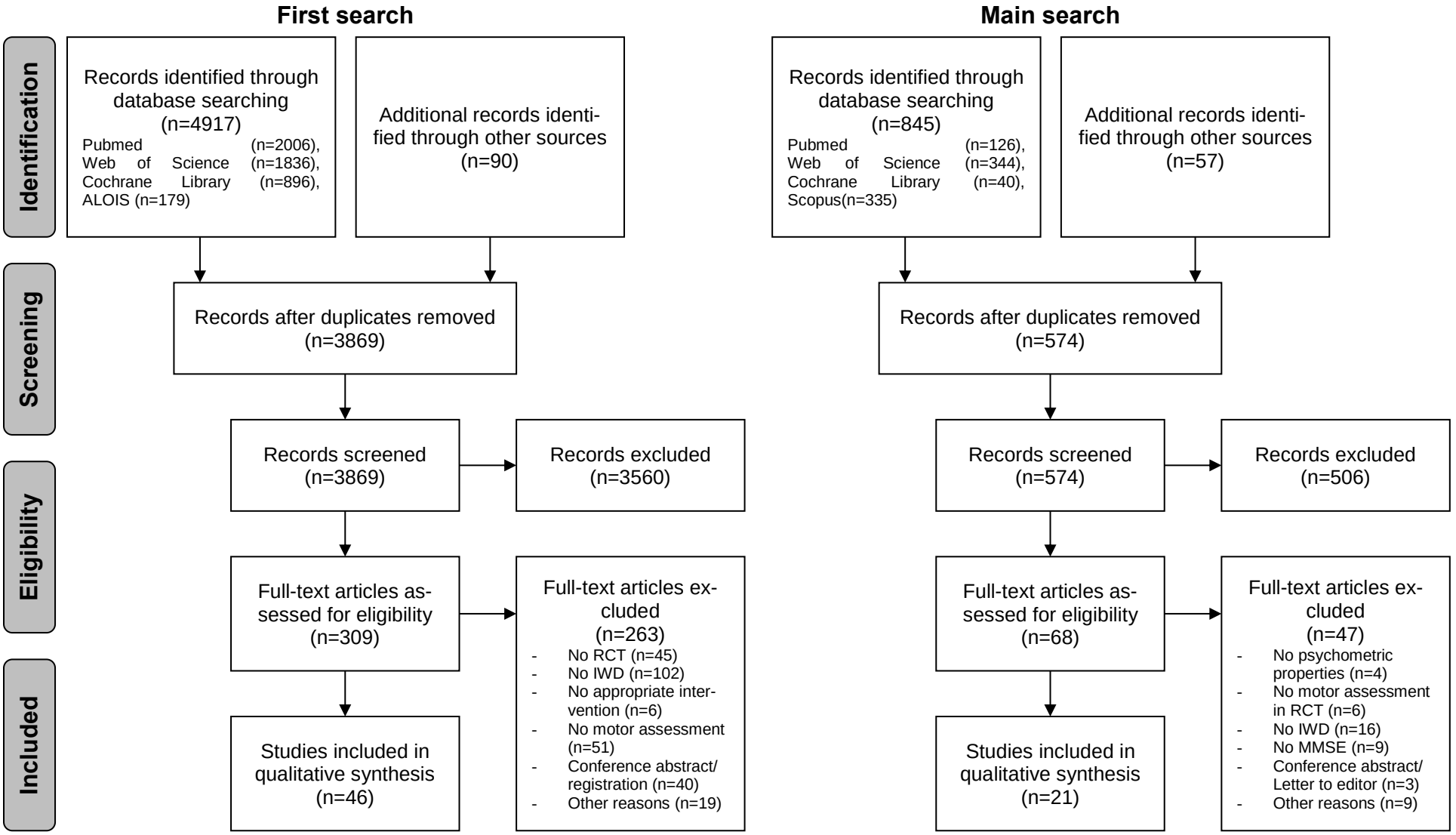


Figure 4. Flow of information (IWD: individuals with dementia, MMSE: Mini-Mental State Examination, n: number, RCT: randomised controlled trial).

5.2.4.2 Motor assessments applied in previous randomised controlled trials

Previous RCT with IWD utilised 57 different motor assessments to determine balance, mobility and gait, strength, endurance, flexibility, and functional performance. Psychometric properties of 28 of these assessments were investigated in IWD. Table 13 contains a short description of all identified motor assessments with available psychometric property studies (see Additional file 9 for motor assessments identified during first search without available information on psychometric properties).

Table 13. Description, frequency of use, and effect sizes of motor assessments applied in previous randomised controlled trials

Motor assessment	Description	Frequency of use	Time*group interaction effect size
	Balance		
FICSIT-4 (Rossiter-Fornoff et al., 1995)	*Task*: performing four different stances with eyes open for ten seconds: (a) feet together, (b) semi-tandem, (c) tandem, (d) single-leg *Measurement*: score [0-5], which rates performance according to ability to maintain stances	1 RCT (n=109) (Bossers et al., 2015)	-
Modified Clinical Test of Sensory Interaction of Balance (Suttanon et al., 2011)	*Task*: standing on a platform (NeuroCom Balance Master) as quietly as possible for ten seconds under four sensory conditions: eyes open and closed standing on firm surface and foam *Measurements*: sway velocity [deg/s], composite score for all conditions	1 RCT (n=40) (Suttanon et al., 2013)	-
Limits of Stability (Suttanon et al., 2011)	*Task*: standing on NeuroCom Balance Master and moving cursor from centre box directly to eight target boxes as fast and as close as possible by shifting weight *Measurements*: reaction time [s], movement velocity [deg/s], maximum excursion [%], directional control [%], summary composite score	1 RCT (n=40) (Suttanon et al., 2013)	-
Physiomat-Trail-Making Task (Wiloth, Lemke, Werner, & Hauer, 2016)	*Task*: standing on Physiomat and connecting digits by shifting weight *Measurements*: total duration [s], accuracy of sway path [digits/ms]	1 RCT (n=84) (Wiloth et al., 2018)	-

Physiomat-Follow-The-Ball Task (Wiloth et al., 2016)	*Task*: standing on Physiomat and moving cursor from centre of screen directly to targets as fast as possible by shifting weight *Measurements*: total duration [s], accuracy of sway path [digits/ms]	1 RCT (n=84) (Wiloth et al., 2018)	-
FR (Duncan et al., 1990)	*Task*: standing next to a wall, holding one arm parallel to a metre stick attached to the wall at shoulder height, and reaching forward as far as possible without losing balance or changing foot position *Measurement*: distance from starting to end position [cm]	5 RCT (n=204) (Arcoverde et al., 2014; Miu et al., 2008; Netz et al., 2007; Suttanon et al., 2013; Vreugdenhil et al., 2012)	Small to large [c]
Hill Step Test (Hill, 1996)	*Task*: stepping one foot onto a block and returning it to the floor as quickly as possible for fifteen seconds *Measurement*: number of repetitions	2 RCT (n=54) (Suttanon et al., 2013; Wesson et al., 2013)	-
Step Quick Turn Test (Suttanon et al., 2011)	*Task*: taking two steps forward on NeuroCom Balance Master, quickly turning, and returning to starting point *Measurements*: turn time [s], turn sway [deg/s]	1 RCT (n=40) (Suttanon et al., 2013)	-
Figure of Eight Test (Johansson & Jarnlo, 2009)	*Task*: walking a lap of a standard figure-eight trajectory as quickly and accurately as possible *Measurements*: walking speed [m/s], number of oversteps	1 RCT (n=109) (Bossers et al., 2015)	-
GMWT (Bossers, van der Woude et al., 2014)	*Task*: walking over a meandering curved line as quickly and accurately as possible *Measurements*: walking speed [m/s], number of oversteps	1 RCT (n=109) (Bossers et al., 2015)	-
BBS (Berg, 1989)	*Task*: 14-item functional balance assessment with simple everyday tasks (reaching, bending, transferring, standing, and rising), which are graded on a five-point ordinal scale (0 to 4) *Measurement*: score [0-56]	11 RCT (n=648) (Arcoverde et al., 2014; Burgener et al., 2008; Christofoletti et al., 2008; M.-J. Kim et al., 2016; Lam, Liao et al., 2018; Miu et al., 2008; Padala et al., 2012; Padala et al., 2017; Telenius et al., 2015a; Toots et al., 2016; Yoon et al., 2013)	Small to large [c/r]
Modified BBS (Berg, 1989)	*Task*: abbreviated version of the original 14-item BBS, excluding three items (chair-to-chair transfer, forward reach with outstretched arm, and alternate stepping on-off stool) *Measurement*: score [0-44]	1 RCT (n=23) (Dawson et al., 2019)	-

POMA (Tinetti, 1986)	*Task*: scale with two parts, assessing balance (B) and gait (G) (B) sitting balance, rising from a chair and sitting down, standing balance (with eyes open and closed), and turning balance (G) gait initiation, step length and height, symmetry, continuity, path direction, and trunk sway *Measurements*: total score [0-28], balance score [0-16], gait score [0-12]	7 RCT (n=300) (Francese et al., 1997; Hauer et al., 2012; Hauer et al., 2017; Kovács et al., 2013; Lam, Liao et al., 2018; Padala et al., 2012; Santana-Sosa et al., 2008)	No to large [c/r]
	Mobility and gait		
TUG (Podsiadlo & Richardson, 1991)	*Task*: standing up from a chair, walking three metres, turning around, walking back to chair, and sitting down *Measurements*: time [s], number of steps	16 RCT (n=1001) (Aguiar et al., 2014; Arcoverde et al., 2014; Bossers et al., 2015; Cancela et al., 2016; Christofoletti et al., 2008; Hauer et al., 2012; Kampragkou et al., 2017; Kovács et al., 2013; Lam, Liao et al., 2018; Netz et al., 2007; Padala et al., 2012; Sobol et al., 2016; Suttanon et al., 2013; Toulotte et al., 2003; Vreugdenhil et al., 2012; Yoon et al., 2013)	No to large [c/r]
Cognitive TUG (Shumway-Cook et al., 2000)	*Task*: TUG with additional cognitive task (counting backwards by threes/evoke names of animals) *Measurement*: time [s]	2 RCT (n=60) (Arcoverde et al., 2014; Suttanon et al., 2013)	-
Manual TUG (Lundin-Olsson et al., 1998; Shumway-Cook et al., 2000)	*Task*: TUG with additional manual task (carrying a glass of water) *Measurement*: time [s]	1 RCT (n=40) (Suttanon et al., 2013)	-
6m WT (Guralnik, Seeman et al., 1994)	*Task*: walking six metres with comfortable pace *Measurements*: walking speed [m/s], step length [m]	3 RCT (n=379) (Bossers et al., 2015; Rolland et al., 2007; Telenius et al., 2015a)	-
4m WT (Guralnik, Seeman et al., 1994)	*Task*: walking four metres with comfortable pace *Measurement*: walking speed [m/s]	2 RCT (n=244) (Souto Barreto et al., 2017; Toots et al., 2017)	Small [c/r]

Instrumented gait analysis (Kressig & Beauchet, 2006)	*Task*: walking with comfortable/fast pace over an electronic walkway (GAITRite, Bessou locometer, NeuroCom Balance Master) *Measurements*: walking speed [cm/s, m/s], cadence [steps/min], stride/step length [cm, m], stride time [s], double/single support [% of stride time], double limb support time [s], step width [cm], step time variability [CV], Walk-Ratio [step length/cadence]	6 RCT (n=370) (Hauer et al., 2012; Kemoun et al., 2010; Pedrinolla et al., 2018; Schwenk, Dutzi et al., 2014; Schwenk, Zieschang et al., 2014; Suttanon et al., 2013)	Small to large [c/r]
	Strength		
5x STS (Csuka & McCarty, 1985)	*Task*: performing five repetitions of the STS task without upper extremity assistance *Measurement*: time [s]	7 RCT (n=358) (Hauer et al., 2012; Hauer et al., 2017; Lam, Liao et al., 2018; Netz et al., 2007; Schwenk, Dutzi et al., 2014; Steinberg et al., 2009; Suttanon et al., 2013)	No to large [c/r]
STS on NeuroCom Balance Master (Suttanon et al., 2011)	*Task*: standing up from a seated position without upper extremity assistance *Measurements*: rising index [% of body weight], centre of gravity sway velocity [deg/s]	1 RCT (n=40) (Suttanon et al., 2013)	-
ACSID (Werner, Wiloth, Lemke, Kronbach, & Hauer, 2018)	*Task*: performing five repetitions of the STS task without upper extremity assistance while motor and cognitive aspects of movement process are qualitatively rated *Measurements*: total score [0-10], sub scores 'recall and initiation' [0-5], 'effective performance' [0-5]	1 RCT (n=77) (Werner et al., 2017)	Large [c/r]
30s CST (Blankevoort et al., 2013; Jones et al., 1999)	*Task*: performing as many repetitions of STS task as possible in 30 seconds *Modified version*: use of upper extremity assistance is allowed *Measurement*: number of repetitions	5 RCT (n=408) (Arcoverde et al., 2014; Dawson et al., 2019; Santana-Sosa et al., 2008; Sobol et al., 2016; Telenius et al., 2015a) Modified: 1 RCT (n=109) (Bossers et al., 2015)	Large [c/r]
Handgrip dynamometer (Thomas & Hageman, 1999)	*Task*: putting maximum force on a dynamometer *Measurement*: maximum handgrip strength [KPa, kg]	3 RCT (n=263) (Hauer et al., 2012; M.-J. Kim et al., 2016; Schwenk, Dutzi et al., 2014)	No [r]

Maximum isometric strength assessed with dynamometers (Verkerke et al., 2003)	*Task*: pushing as hard as possible against a dynamometer after adopting a standardised position *Measurements*: maximum strength [N] and integral over time [Ns] for knee extension, knee flexion, and ankle flexion	2 RCT (n=216) (Bossers et al., 2015; Hauer et al., 2012)	-
	Endurance		
6min WT (Enright, 2003)	*Task*: walking for six minutes with comfortable pace *Measurement*: distance [m, ft]	5 RCT (n=359) (Bossers et al., 2015; Miu et al., 2008; Roach et al., 2011; Tappen et al., 2000; Venturelli et al., 2011)	-
	Functional performance		
SPPB (Guralnik, Simonsick et al., 1994)	*Task*: three subtests including standing balance (tandem, semi-tandem, and side-by-side stands), walking speed over an 8-foot walking course, and 5x STS *Measurement*: score [0-12]	3 RCT (n=313) (Hauer et al., 2017; Pitkälä, Pöysti et al., 2013; Souto Barreto et al., 2017)	Small to medium [c/r]
E-ADL Test (Graessel et al., 2009; Luttenberger et al., 2012)	*Task*: five items (pouring a drink, spreading butter on a sandwich and cutting the sandwich, open a small cupboard with a key, washing and drying hands, and tying a bow on a small wrapped present), which are rated according to correctly performed substeps (0-6 points) *Measurement*: score [0-30]	2 RCT (n=192) (Bossers et al., 2016; Henskens et al., 2018)	-

4m WT: 4-metre walk test, 5x STS: Five Times Sit-to-Stand Test, 6m WT: 6-metre walk test, 6min WT: 6-minute walk test, 30s CST: 30-second chair stand test, ACSID: Assessment of Compensatory Sit-to-Stand Maneuvers in People With Dementia, BBS: Berg Balance Scale, E-ADL Test: Erlangen Test of Activities of Daily Living, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, FR: Functional Reach Test, GMWT: Groningen Meander Walking Test, n: number of analysed participants, POMA: Performance Oriented Mobility Assessment, RCT: randomised controlled trial/s, SPPB: Short Physical Performance Battery, STS: Sit-to-Stand, TUG: Timed Up & Go Test

[c] calculated effect size, [r] effect size provided of randomised controlled trial

5.2.4.3 Psychometric properties

Seventeen of twenty-one studies examining psychometric properties focused on inter-rater and/or test-retest reliability. Herein, they determined consistency among different evaluators simultaneously rating the same participant, and between repeated measurements, respectively (Carter et al., 2013). Investigations assessing content, construct, and criterion validity, internal consistency, and intra-rater reliability were rare. Thus, we only summarised results and did not derive conclusions.

5.2.4.3.1 Summary for content, construct, and criterion validity, internal consistency, and intra-rater reliability[9]

The systematic search did not identify any investigation examining content validity. Based on hypotheses testing or revealing known group differences, construct validity was suggested for Physiomat assessments, the Erlangen Test of Activities of Daily Living (E-ADL Test), and knee extensor strength assessed with dynamometers (Graessel et al., 2009; Luttenberger et al., 2012; Suzuki et al., 2009; Wiloth et al., 2016). Seven investigations include information on criterion validity (concurrent and predictive validity), correlation with, or prediction of external criteria. For the E-ADL Test, criterion related validity was determined based on the relation between achieved scores and level of care (Luttenberger et al., 2012). Concurrent validity with spatiotemporal gait parameters or 2D-video motion analysis was established for a modified BBS, Short Physical Performance Battery (SPPB), and Assessment of Compensatory Sit-to-Stand Maneuvers in People With Dementia (ACSID) (McGough et al., 2013; Werner et al., 2018). Moreover, both the SPPB and 6-minute walk test (6min WT) significantly correlated with peak oxygen consumption (assessed with a cycle ergometer test), suggesting that these assessments are useful in identifying individuals with low aerobic capacity (Bronas et al., 2017). Furthermore, knee extensor strength was found to be a significant predictor for several activities of daily living, gait, and sit-to-stand (STS) performance (Suzuki et al., 2009; Suzuki et al., 2012). No predictive validity concerning future falls could be observed for Timed Up & Go Test (TUG), Performance Oriented Mobility Assessment (POMA), and Five Times Sit-to-Stand Test (5x STS) (Schwenk, Hauer et al., 2014).

Considering internal consistency, three studies observed Cronbach's α between 0.37 and 0.77 for E-ADL Test (Graessel et al., 2009; Luttenberger et al., 2012) and 0.95 for BBS (Telenius et al., 2015b). Furthermore, one study examining ACSID total score determined intra-rater reliability based on ICC ranging between 0.72 and 0.90 (Werner et al., 2018).

[9] This summary utilises psychometric property terms indicated in original studies. These terms have not been consistently used throughout the literature and should have been adapted according to the COSMIN checklist Mokkink, Prinsen et al. (2018).

5.2.4.3.2 Inter-rater reliability (relative and absolute reliability)

Five studies assessed inter-rater reliability of nine assessments. ICC ranged from 0.72 to 1.00 and $MDC_{95\%}$ included values between 0.0 % and 98.0 % (H.-S. Lee et al., 2017; Muir-Hunter et al., 2015; Tappen et al., 1997; Telenius et al., 2015b; Werner et al., 2018). Accordingly, all assessments reached sufficient relative inter-rater reliability. Quality of evidence for relative inter-rater reliability was high for BBS, moderate for TUG, and low or very low for all other assessments. Grading $MDC_{95\%}$, TUG and 6-metre walk test (6m WT) showed sufficient absolute inter-rater reliability, while it was insufficient/unacceptable for 4-metre walk test (4m WT), and indeterminate for all other assessments. Quality of evidence for absolute inter-rater reliability was low for 6m WT and 30-second chair stand test (30s CST), and moderate for all remaining assessments (see Table 14).

Table 14. Relative and absolute inter-rater reliability

	Variable	Study	Relative inter-rater reliability							Absolute inter-rater reliability						
			ICC	Rating	Risk of bias	Inconsistency	Imprecision	Indirectness	Quality of evidence	MDC 95%	Rating	Risk of bias	Inconsistency	Imprecision	Indirectness	Quality of evidence
Balance																
FR	Distance	1 study of adequate quality (n=15) (Muir-Hunter et al., 2015)	0.79	+	Serious	No	n<50	No	Very low	Not assessed						
GMWT	Time	1 study of adequate/very good quality (n=53) (H.-S. Lee et al., 2017)	0.99	+	Serious	No	n=50-100	No	Low	14.5%	?	No	No	n=50-100	No	Moderate
	Number of oversteps	1 study of adequate/very good quality (n=53) (H.-S. Lee et al., 2017)	0.99	+	Serious	No	n=50-100	No	Low	17.1%	?	No	No	n=50-100	No	Moderate
BBS	Score	3/2 studies of adequate/very good quality (n=101/86) (H.-S. Lee et al., 2017; Muir-Hunter et al., 2015; Telenius et al., 2015b)	0.72 - 0.99	+	No	No	No	No	High	5.9-7.1%	?	No	No	n=50-100	No	Moderate
Mobility and gait																
TUG	Time	2 studies/1 study of adequate/very good quality (n=68/53) (H.-S. Lee et al., 2017; Muir-Hunter et al., 2015)	0.98 - 0.99	+	No	No	n=50-100	No	Moderate	7.9%	+[b]	No	No	n=50-100	No	Moderate
6m WT	Walking speed	1 study of adequate/very good quality (n=33) (Telenius et al., 2015b)	0.97	+	Serious	No	n<50	No	Very low	15.7%	+[c]	No	No	n<50	No	Low
4m WT	Time	1 study of adequate/very good quality (n=53) (H.-S. Lee et al., 2017)	0.82	+	Serious	No	n=50-100	No	Low	98.0%	-[c]/↓	No	No	n=50-100	No	Moderate
Strength																
ACSID	Score	1 study of very good quality (n=94) (Werner et al., 2018)	0.85	+	No	No	n=50-100	No	Moderate	Not assessed						
30s CST	Repetitions	1 study of adequate/very good quality (n=33) (Telenius et al., 2015b)	1.00	+	Serious	No	n<50	No	Very low	0.0%	?	No	No	n<50	No	Low

Endurance										
6min WT	Distance	1 study of adequate quality (n=33)[a] (Tappen et al., 1997)	0.97 - 0.99	+	Seri-ous	No	n<50	No	Very low	Not assessed
	Walking speed	1 study of adequate quality (n=33)[a] (Tappen et al., 1997)	0.96 - 0.98	+	Seri-ous	No	n<50	No	Very low	Not assessed

4m WT: 4-metre walk test, 6m WT: 6-metre walk test, 6min WT: 6-minute walk test, 30s CST: 30-second chair stand test, ACSID: Assessment of Compensatory Sit-to-Stand Maneuvers in People With Dementia, BBS: Berg Balance Scale, FR: Functional Reach Test, GMWT: Groningen Meander Walking Test, ICC: intraclass correlation coefficient, $MDC_{95\%}$: percentage minimal detectable changes at 95 % confidence interval, n: total number of participants, TUG: Timed Up & Go Test
Rating according to COSMIN criteria for good measurement properties: +=sufficient, -=insufficient, ?=indeterminate, ↓=unacceptable absolute inter-rater reliability
[a] inter-rater reliability was determined on 2 times of measurement, [b] minimal important change (TUG)=10.1 s (Blankevoort et al., 2013; van Iersel, Munneke, Esselink, Benraad, & Olde Rikkert, 2008), [c] minimal important change (walking speed)=0.21 m/s (Blankevoort et al., 2013; van Iersel et al., 2008)

Regarding balance assessments, ICC were higher for Groningen Meander Walking Test (GMWT) and BBS than for Functional Reach Test (FR). Furthermore, $MDC_{95\%}$ were lower for BBS compared to GMWT. Focusing on GMWT, time measurement showed lower $MDC_{95\%}$ than number of oversteps. For mobility and gait, ICC increased and $MDC_{95\%}$ decreased from 4m WT, through 6m WT, to TUG. Considering strength assessments, ICC were higher for 30s CST counting repetitions than for ACSID rating STS performance, while $MDC_{95\%}$ was only determined for 30s CST. Since ICC was only assessed for 6min WT, a comparison of inter-rater reliability of endurance assessments was not possible (see Table 14).

5.2.4.3.3 Test-retest reliability (relative and absolute reliability)

Fifteen studies investigated test-retest reliability considering 24 assessments. ICC ranged between 0.02 and 0.99 and $MDC_{95\%}$ varied from 6.8 % to 225.7 % (Alencar et al., 2012; Blankevoort et al., 2013; Bossers, van der Woude et al., 2014; Graessel et al., 2009; H.-S. Lee et al., 2017; McGough et al., 2013; Muir-Hunter et al., 2015; Ries et al., 2009; Suttanon et al., 2011; Suzuki et al., 2009; Tappen et al., 1997; Thomas & Hageman, 1999; Wiloth et al., 2016; Wittwer et al., 2008; Wittwer et al., 2013) (see Table 15).

Table 15. Relative and absolute test-retest reliability

	Variable	Study	Relative test-retest reliability							Absolute test-retest reliability						
			ICC	Rating	Risk of bias	Inconsistency	Imprecision	Indirectness	Quality of evidence	MDC 95%	Rating	Risk of bias	Inconsistency	Imprecision	Indirectness	Quality of evidence
Balance																
FICSIT-4	Score	1 study of adequate quality (n=58)[a] (Blankevoort et al., 2013)	0.79 - 0.82	+	Serious	No	n=50-100	No	Low	58.9-71.1%	↓	Serious	No	n=50-100	No	Low
Modified Clinical Test of Sensory Interaction of Balance	Sway velocity	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.91	+	Serious	No	n<50	No	Very low	36.5%	↓	Serious	No	n<50	No	Very low
Limits of Stability	Reaction time	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.52	-	Serious	No	n<50	No	Very low	38.0%	↓	Serious	No	n<50	No	Very low
	Movement velocity	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.48	-	Serious	No	n<50	No	Very low	38.9%	↓	Serious	No	n<50	No	Very low
	Maximum excursion	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.68	-	Serious	No	n<50	No	Very low	15.9%	?	Serious	No	n<50	No	Very low
	Directional control	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.71	+	Serious	No	n<50	No	Very low	21.8%	?	Serious	No	n<50	No	Very low
Physiomat-Trail-Making Task	Score	1 study of adequate quality (n=74) (Wiloth et al., 2016)	0.90	+	Serious	No	n=50-100	No	Low				Not assessed			
	Sway Path	1 study of adequate quality (n=47-73)[b] (Wiloth et al., 2016)	0.47 - 0.82	+/- depending on condition	Serious	No	n=50-100	No	Low				Not assessed			
	Time	1 study of adequate quality (n=47-73)[b] (Wiloth et al., 2016)	0.55 - 0.83	+/- depending on condition	Serious	No	n=50-100	No	Low				Not assessed			
Physiomat-Follow-The-Ball Task	Sway Path	1 study of adequate quality (n=73) (Wiloth et al., 2016)	0.84	+	Serious	No	n=50-100	No	Low				Not assessed			

	Time	1 study of adequate quality (n=73) (Wiloth et al., 2016)	0.79	+	Serious	No	n=50-100	No	Low	Not assessed						
FR	Distance	2 studies of adequate quality (n=29) (Muir-Hunter et al., 2015; Suttanon et al., 2011)	0.81 - 0.84	+	No	No	n<50	No	Low	15.4-68.9%	?/↓	No	Yes	n<50	No	Not assigned (inconsistency)
Hill Step Test	Number of steps	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.87	+	Serious	No	n<50	No	Very low	26.2%	?	Serious	No	n<50	No	Very low
Step Quick Turn Test	Time	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.55	-	Serious	No	n<50	No	Very low	38.1%	↓	Serious	No	n<50	No	Very low
	Sway	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.64	-	Serious	No	n<50	No	Very low	29.7%	?	Serious	No	n<50	No	Very low
Figure of Eight Test	Time	1 study of adequate quality (n=46)[a] (Blankevoort et al., 2013)	0.85 - 0.94	+	Serious	No	n<50	No	Very low	36.9-37.9%	↓	Serious	No	n<50	No	Very low
GMWT	Time	2 studies of adequate quality (n=95)[a] (Bossers, van der Woude et al., 2014; H.-S. Lee et al., 2017)	0.93 - 0.99	+	No	No	n=50-100	No	Moderate	19.6-31.2%	?/↓	No	No	n=50-100	No	Moderate
	Number of oversteps	2 studies of adequate quality (n=95)[a] (Bossers, van der Woude et al., 2014; H.-S. Lee et al., 2017)	0.57 - 0.96	?	No	Yes	n=50-100	No	Not assigned (inconsistency)	33.3-225.7 %	↓	No	Yes	n=50-100	No	Not assigned (inconsistency)
BBS	Score	2 studies of adequate quality (n=68) (H.-S. Lee et al., 2017; Muir-Hunter et al., 2015)	0.95 - 0.99	+	No	No	n=50-100	No	Moderate	10.2-38.6%	?/↓	No	No	n=50-100	No	Moderate
Mobility and gait																
TUG	Time	6/5 studies of adequate quality (n=200/191)[a] (Blankevoort et al., 2013; H.-S. Lee et al., 2017; Muir-Hunter et al., 2015; Ries et al., 2009; Suttanon et al., 2011; Thomas & Hageman, 1999)	0.72 - 0.99	+	No	No	No	No	High	15.8-39.6%	+[h]/↓	No	No	No	No	High

Cognitive TUG	Time	1 study of adequate quality (n=10) (Suttanon et al., 2011)	0.51	-	Serious	No	n<50	No	Very low	36.2%	+[h]/↓	Serious	No	n<50	No	Very low
Manual TUG	Time	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.70	+	Serious	No	n<50	No	Very low	26.7%	+[h]	Serious	No	n<50	No	Very low
6m WT	Walking speed	1 study of adequate quality (n=58)[a] (Blankevoort et al., 2013)	0.83 - 0.89	+	Serious	No	n=50-100	No	Low	31.6-41.5%	-[i]/↓	Serious	No	n=50-100	No	Low
	Time	1 study of adequate quality (n=9-10)[b] (Thomas & Hageman, 1999)	0.92 - 0.95	+	Serious	No	n<50	No	Very low	Not assessed						
	Number of steps	1 study of adequate quality (n=9-10)[b] (Thomas & Hageman, 1999)	0.80 - 0.90	+	Serious	No	n<50	No	Very low	Not assessed						
4m WT	Time	1 study of adequate quality (n=53) (H.-S. Lee et al., 2017)	0.85	+	Serious	No	n=50-100	No	Low	84.3%	-[i]/↓	Serious	No	n=50-100	No	Low
Instrumented gait analysis	Walking speed	4/3 studies of adequate quality (n=93/85)[a, d, e] (McGough et al., 2013; Ries et al., 2009; Suttanon et al., 2011; Wittwer et al., 2008)	0.50 - 0.98	+ (except for NeuroCom Balance Master)	No	No	n=50-100	No	Moderate	10.2-48.3%	+[i]/↓	No	No	n=50-100	No	Moderate
	Step length	2 studies of adequate quality (n=34)[a, d, e] (Suttanon et al., 2011; Wittwer et al., 2008)	0.75 - 0.98	+	No	No	n<50	No	Low	7.0-35.6%	?/↓	No	No	n<50	No	Low
	Step width	2 studies of adequate quality (n=34) [a, d, e] (Suttanon et al., 2011; Wittwer et al., 2008)	0.89 - 0.95	+	No	No	n<50	No	Low	20.0-24.7%	?	No	No	n<50	No	Low
	Stride length	2 studies/1 study of adequate quality (n=28/20)[e] (McGough et al., 2013; Wittwer et al., 2008)	0.97 - 0.98	+	No	No	n<50	No	Low	6.8-8.5%	?	Serious	No	n<50	No	Very low
	Cadence	2 studies/1 study of adequate quality (n=28/20)[e] (McGough et al., 2013; Wittwer et al., 2008)	0.88 - 0.91	+	No	No	n<50	No	Low	7.1-7.5%	?	Serious	No	n<50	No	Very low

	Swing time	2 studies/1 study of adequate quality (n=28/20)[e] (McGough et al., 2013; Wittwer et al., 2008)	0.89 - 0.96	+	No	No	n<50	No	Low	7.0-7.1%	?	Serious	No	n<50	No	Very low
	Stance time	1 study of adequate quality (n=20)[e] (Wittwer et al., 2008)	0.70 - 0.73	+	Serious	No	n<50	No	Very Low	8.6-8.7%	?	Serious	No	n<50	No	Very low
	Toe in/out angle	1 study of adequate quality (n=20)[e] (Wittwer et al., 2008)	0.91 - 0.93	+	Serious	No	n<50	No	Very Low	28.2-33.5%	?/↓	Serious	No	n<50	No	Very low
	Walking speed variability	1 study of adequate quality (n=16) (Wittwer et al., 2013)	0.66	-	Serious	No	n<50	No	Very Low	77.8%	↓	Serious	No	n<50	No	Very low
	Stride length variability	1 study of adequate quality (n=16) (Wittwer et al., 2013)	0.80	+	Serious	No	n<50	No	Very Low	71.7%	↓	Serious	No	n<50	No	Very low
	Stride width variability	1 study of adequate quality (n=16) (Wittwer et al., 2013)	0.83	+	Serious	No	n<50	No	Very Low	46.9%	↓	Serious	No	n<50	No	Very low
	Cadence variability	1 study of adequate quality (n=16) (Wittwer et al., 2013)	0.65	-	Serious	No	n<50	No	Very Low	41.4%	↓	Serious	No	n<50	No	Very low
Strength																
5x STS	Time	2 studies/1 study of adequate quality (n=24/14) (Suttanon et al., 2011; Thomas & Hageman, 1999)	0.80 - 0.94	+	No	No	n<50	No	Low	29.9%	?	Serious	No	n<50	No	Very low
STS on NeuroCom Balance Master	Rising Index	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.95	+	Serious	No	n<50	No	Very low	21.8%	?	Serious	No	n<50	No	Very low
	COG sway velocity	1 study of adequate quality (n=14) (Suttanon et al., 2011)	0.02	-	Serious	No	n<50	No	Very low	80.2%	↓	Serious	No	n<50	No	Very low
Modified 30s CST	Repetitions	1 study of adequate quality (n=52)[a] (Blankevoort et al., 2013)	0.79 - 0.88	+	Serious	No	n=50-100	No	Low	33.2-45.7%	↓	Serious	No	n=50-100	No	Low

Handgrip dynamometer	Force	3 studies/1 study of adequate quality (n=143/57)[a] (Alencar et al., 2012; Blankevoort et al., 2013; Thomas & Hageman, 1999)	0.42 - 0.98	+ (except for severe dementia)	No	No	No	No	High	34.9-36.8%	↓	Serious	No	n=50-100	No	Low
Maximum isometric strength assessed with dynamometers	Peak force	1 studies of adequate quality (n=11-12)[f] (Thomas & Hageman, 1999)	0.63 - 0.71	?	Serious	Yes	n<50	No	Not assigned (inconsistency)	Not assessed						
	(Normalised) torque	1 studies of adequate quality (n=60)[a] (Suzuki et al., 2009)	0.95 - 0.98	+	Serious	No	n=50-100	No	Low	Not assessed						
Endurance																
6min WT	Distance	2 studies/1 study of adequate quality (n=84/51)[a, c] (Ries et al., 2009; Tappen et al., 1997)	0.76 - 0.98	+	No	No	n=50-100	No	Moderate	21.2-28.9%	?	Serious	No	n=50-100	No	Low
	Walking speed	1 study of adequate quality (n=33)[c] (Tappen et al., 1997)	0.75 - 0.89	+	Serious	No	n<50	No	Very Low	Not assessed						
Functional performance																
E-ADL Test	Score	1 study of doubtful quality (n=42) (Graessel et al., 2009)	r= 0.73 [g]	?	Very serious	No	n<50	No	Very Low	Not assessed						

4m WT: 4-metre walk test, 5x STS: Five Times Sit-to-Stand Test, 6m WT: 6-metre walk test, 6min WT: 6-minute walk test, 30s CST: 30-second chair stand test, BBS: Berg Balance Scale, COG: centre of gravity, E-ADL Test: Erlangen Test of Activities of Daily Living, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, FR: Functional Reach Test, GMWT: Groningen Meander Walking Test, ICC: intraclass correlation coefficient, $MDC_{95\%}$: percentage minimal detectable changes at 95 % confidence interval, n: total number of participants, STS: Sit-to-Stand, TUG: Timed Up & Go Test

Rating according to COSMIN criteria for good measurement properties: +=sufficient, -=insufficient, ?=indeterminate, ↓=unacceptable absolute test-retest reliability

[a] test-retest reliability was assessed for different subgroups, [b] test-retest reliability was assessed for different conditions, [c] test-retest reliability was assessed for 2 different raters and 2 different between-test intervals, [d] test-retest reliability was assessed with 2 different devices, [e] test-retest reliability was assessed with 2 analysis sets, [f] test-retest reliability was assessed for 3 muscle groups, [g] Spearman's rank correlation coefficient, [h] minimal important change (TUG)=10.1 s (Blankevoort et al., 2013; van Iersel et al., 2008), [i] minimal important change (walking speed)=0.21 m/s (Blankevoort et al., 2013; van Iersel et al., 2008)

Most studies focused on between-day test-retest reliability, while some studies examined within-day and within-session test-retest reliability. Comparing these studies, ICC increased and $MDC_{95\%}$ decreased, respectively, from between-day (ICC=0.02-0.99, $MDC_{95\%}$=6.8-225.7 %; Alencar et al., 2012; Blankevoort et al., 2013; Bossers, van der Woude et al., 2014; H.-S. Lee et al., 2017; Muir-Hunter et al., 2015; Suttanon et al., 2011; Tappen et al., 1997; Thomas & Hageman, 1999; Wiloth et al., 2016; Wittwer et al., 2008; Wittwer et al., 2013), through within-day (ICC=0.79-0.99, $MDC_{95\%}$=21.1-30.0 %; McGough et al., 2013; Ries et al., 2009; Tappen et al., 1997), to within-session test-retest reliability (ICC=0.95-0.98; Suzuki et al., 2009).

5.2.4.3.3.1 Balance

Six investigations assessing test-retest reliability of eleven balance assessments determined ICC and $MDC_{95\%}$ ranging between 0.32-0.99 and 10.2-225.7 %, respectively (Blankevoort et al., 2013; Bossers, van der Woude et al., 2014; H.-S. Lee et al., 2017; Muir-Hunter et al., 2015; Suttanon et al., 2011; Wiloth et al., 2016). Relative test-retest reliability was sufficient for all balance assessments except for Limits of Stability, Step Quick Turn Test, and simple condition of Physiomat-Trail-Making Task. However, quality of evidence for relative test-retest reliability was low or very low for most assessments. Only GMWT (time) and BBS reached moderate quality of evidence. Absolute test-retest reliability for balance assessments was indeterminate or unacceptable with moderate to very low quality of evidence (see Table 15).

GMWT (time) and BBS showed the highest ICC, while we could not observe a clear tendency for $MDC_{95\%}$. Comparing different outcomes of GMWT, ICC were higher and $MDC_{95\%}$ were lower for time than for number of oversteps (see Table 15).

5.2.4.3.3.2 Mobility and gait

Nine studies investigated test-retest reliability of six mobility and gait assessments. They reported ICC between 0.50 and 0.99 and $MDC_{95\%}$ from 6.8 % to 84.3 % (Blankevoort et al., 2013; H.-S. Lee et al., 2017; McGough et al., 2013; Muir-Hunter et al., 2015; Ries et al., 2009; Suttanon et al., 2011; Thomas & Hageman, 1999; Wittwer et al., 2008; Wittwer et al., 2013). Relative test-retest reliability was sufficient for TUG, manual TUG, 6m WT, 4m WT, and instrumented gait analysis (except for cadence variability, walking speed variability, and walking speed assessed with NeuroCom Balance Master), while it was insufficient for cognitive TUG. Quality of evidence for relative

test-retest reliability was high for TUG, moderate to very low for instrumented gait analysis, and low or very low for all other assessments. Absolute test-retest reliability was indeterminate for spatiotemporal gait parameters, insufficient/unacceptable for variability gait parameters, 4m WT, and 6m WT, and sufficient for manual TUG. For TUG, cognitive TUG, and walking speed assessed with instrumented gait analysis, absolute test-retest reliability was sufficient according to COSMIN criteria but unacceptable when applying $MDC_{95\%}$ limit of 30 %. Except for TUG and walking speed assessed with instrumented gait analysis (high/moderate quality of evidence), quality of evidence for absolute test-retest reliability was low or very low (see Table 15).

Considering up and go tasks, ICC were higher for single than for dual task conditions. Focusing on short distance walk tests (WT), $MDC_{95\%}$ were lower for 6m WT than for 4m WT. Furthermore, the comparison of different gait parameters assessed with instrumented gait analysis, determined lower ICC and higher $MDC_{95\%}$ for variability measures than for spatiotemporal gait parameters. Comparing different assessments to determine short distance walking speed showed higher ICC and lower $MDC_{95\%}$ for instrumented gait analysis (except for NeuroCom Balance Master) than for simple short distance WT (see Table 15).

5.2.4.3.3.3 Strength

Five studies focusing on test-retest reliability of strength assessments reported ICC and $MDC_{95\%}$ ranging between 0.02-0.98 and 21.8 %-80.2 %, respectively (Alencar et al., 2012; Blankevoort et al., 2013; Suttanon et al., 2011; Suzuki et al., 2009; Thomas & Hageman, 1999). Relative test-retest reliability was sufficient for modified 30s CST, 5x STS, handgrip dynamometers (except for severe dementia and one-time measuring), and maximum isometric strength assessed with dynamometers (except for dorsiflexor and iliopsoas muscle strength), while it was insufficient for STS on NeuroCom Balance Master (except for Rising Index). Quality of evidence for relative test-retest reliability was high for handgrip dynamometers and low or very low for all other strength assessments. Absolute test-retest reliability was indeterminate for 5x STS and Rising Index of STS on NeuroCom Balance Master, and unacceptable for modified 30s CST, centre of gravity sway velocity of STS on NeuroCom Balance Master, and handgrip dynamometers. Quality of evidence for absolute test-retest reliability was low or very low for all assessments (see Table 15).

Comparing different STS assessments, ICC for assessments performing only one STS repetition were lower (except for Rising Index) than STS assessments with more repetitions. Moreover, $MDC_{95\%}$ increased from 5x STS, through modified 30s CST, to STS on NeuroCom Balance Master (except for Rising Index) (see Table 15).

5.2.4.3.3.4 Endurance

Considering endurance, test-retest reliability was only determined for 6min WT. Two studies observed ICC between 0.75 and 0.98, while $MDC_{95\%}$ ranged from 21.2 % to 28.9 % (Ries et al., 2009; Tappen et al., 1997). Accordingly, relative test-retest reliability was sufficient with moderate to very low quality of evidence. Absolute test-retest reliability was indeterminate with low quality of evidence (see Table 15).

5.2.4.3.3.5 Functional Performance

Functional performance was rarely assessed. One study focusing on the E-ADL Test did not determine ICC and $MDC_{95\%}$, but found significant correlations for the whole test (r=0.73) and separate items (r=0.35-0.63) (Graessel et al., 2009). Quality of evidence was very low.

5.2.4.3.4 Influence of severity and aetiology of dementia and cueing on test-retest reliability

With respect to severity of dementia, the Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4 (FICSIT-4) and GMWT tend to yield higher ICC and/or lower $MDC_{95\%}$ with less cognitive impairment. In contrast, ICC were slightly higher and/or $MDC_{95\%}$ lower with stronger cognitive impairment for BBS, 6m WT, modified 30s CST, and 5x STS (see Table 16).

Table 16. Subgroup analysis of test-retest reliability considering severity of dementia

	Mild dementia	**Mild to moderate dementia**	**Moderate dementia**	**Severity not reported**
FICSIT-4	MMSE [mean (SD)]: 22.7 (2.1) ICC=0.82 $MDC_{95\%}$=58.9 % (Blankevoort et al., 2013)	MMSE [mean (SD)]: 19.2 (4.4) ICC=0.79 $MDC_{95\%}$=59.4 % (Blankevoort et al., 2013)	MMSE [mean (SD)]: 15.5 (2.4) ICC=0.80 $MDC_{95\%}$=71.1 % (Blankevoort et al., 2013)	

GMWT	MMSE [mean (SD)]: n.r. ICC=0.79-0.96 $MDC_{95\%}$=n.r. (Bossers, van der Woude et al., 2014)	MMSE [mean (SD)]: 17.4 (4.3) ICC=0.63-0.94 $MDC_{95\%}$=31.2-225.7 % (Bossers, van der Woude et al., 2014)	MMSE [mean (SD)]: n.r. ICC=0.57-0.93 $MDC_{95\%}$=n.r. (Bossers, van der Woude et al., 2014)	MMSE [mean (SD)]: 13.8 (5.7) ICC=0.96-0.99 $MDC_{95\%}$=19.6-33.3 % (H.-S. Lee et al., 2017)
BBS		MMSE [mean (SD)]: 20.0 (5.5) ICC=0.95 $MDC_{95\%}$=38.6 % (Muir-Hunter et al., 2015)		MMSE [mean (SD)]: 13.8 (5.7) ICC=0.99 $MDC_{95\%}$=10.2 % (H.-S. Lee et al., 2017)
6m WT	MMSE [mean (SD)]: 22.7 (2.1) ICC=0.83 $MDC_{95\%}$=41.5 % (Blankevoort et al., 2013)	MMSE [mean (SD)]: 19.2 (4.4) ICC=0.86 $MDC_{95\%}$=36.5 % (Blankevoort et al., 2013)	MMSE [mean (SD)]: 15.5 (2.4) ICC=0.89 $MDC_{95\%}$=31.6 % (Blankevoort et al., 2013)	MMSE [mean (SD)]: 16.9 (7.3) ICC=0.80-0.95 $MDC_{95\%}$=n.r. (Thomas & Hageman, 1999)
5x STS	MMSE [mean (SD)]: 21.4 (5.0) ICC=0.80 $MDC_{95\%}$=29.9 % (Suttanon et al., 2011)			MMSE [mean (SD)]: 16.9 (7.3) ICC=0.94 $MDC_{95\%}$=n.r. (Thomas & Hageman, 1999)
Modified 30s CST	MMSE [mean (SD)]: 22.7 (2.1) ICC=0.79 $MDC_{95\%}$=45.7 % (Blankevoort et al., 2013)	MMSE [mean (SD)]: 19.2 (4.4) ICC=0.84 $MDC_{95\%}$=42.5 % (Blankevoort et al., 2013)	MMSE [mean (SD)]: 15.5 (2.4) ICC=0.88 $MDC_{95\%}$=33.2 % (Blankevoort et al., 2013)	

5x STS: Five Times Sit-to-Stand Test, 6m WT: 6-metre walk test, 30s CST: 30-second chair stand test, BBS: Berg Balance Scale, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, GMWT: Groningen Meander Walking Test, ICC: intraclass correlation coefficient, $MDC_{95\%}$: percentage minimal detectable changes at 95 % confidence interval, MMSE: Mini-Mental State Examination, n.r.: not reported, SD: standard deviation

Regarding aetiology of dementia, maximum isometric strength assessed with dynamometers and short distance walking speed (except for instrumented gait analysis with NeuroCom Balance Master) resulted in somewhat higher ICC and/or lower $MDC_{95\%}$ for AD vs. various or not reported types. In contrast, ICC were slightly higher and/or $MDC_{95\%}$ were lower for various or not reported types vs. AD for BBS, TUG (between-day reliability), up and go tasks in general (between-day reliability), 5x STS, and STS tasks in general (except for Rising Index; see Table 17).

Table 17. Subgroup analysis of test-retest reliability considering aetiology of dementia

	Alzheimer's disease	**Various types/not reported**
BBS	ICC=0.95 $MDC_{95\%}$=38.6 % (Muir-Hunter et al., 2015)	ICC=0.99 $MDC_{95\%}$=10.2 % (H.-S. Lee et al., 2017)
TUG (between-day reliability)	ICC=0.72-0.76 ($MDC_{95\%}$=20.3-24.9 %) (Muir-Hunter et al., 2015; Suttanon et al., 2011)	ICC=0.87-0.99 ($MDC_{95\%}$=15.8-39.6 %) (Blankevoort et al., 2013; H.-S. Lee et al., 2017; Thomas & Hageman, 1999)
Up and go tasks (between-day reliability)	ICC=0.51-0.76 ($MDC_{95\%}$=20.3-36.2 %) (Muir-Hunter et al., 2015; Suttanon et al., 2011)	ICC=0.87-0.99 ($MDC_{95\%}$=15.8-39.6 %) (Blankevoort et al., 2013; H.-S. Lee et al., 2017; Thomas & Hageman, 1999)
Short distance walking speed (without NeuroCom Balance Master)	ICC=0.95-0.98 $MDC_{95\%}$=10.2-28.9 % (Ries et al., 2009; Wittwer et al., 2008)	ICC=0.83-0.95 $MDC_{95\%}$=31.6-84.3 % (Blankevoort et al., 2013; H.-S. Lee et al., 2017; McGough et al., 2013)
5x STS	ICC=0.80 $MDC_{95\%}$=29.9 % (Suttanon et al., 2011)	ICC=0.94 $MDC_{95\%}$=n.r. (Thomas & Hageman, 1999)
STS assessments (without Rising Index)	ICC=0.02-0.80 $MDC_{95\%}$=29.9-80.2 % (Suttanon et al., 2011)	ICC=0.79-0.94 $MDC_{95\%}$=33.2-45.7 % (Blankevoort et al., 2013; Thomas & Hageman, 1999)
Maximum isometric strength assessed with dynamometers	ICC=0.95-0.98 $MDC_{95\%}$=n.r. (Suzuki et al., 2009)	ICC=0.63-0.71 $MDC_{95\%}$=n.r. (Thomas & Hageman, 1999)

5x STS: Five Times Sit-to-Stand Test, BBS: Berg Balance Scale, ICC: intraclass correlation coefficient, $MDC_{95\%}$: percentage minimal detectable changes at 95 % confidence interval, n.r.: not reported, STS: Sit-to-Stand, TUG: Timed Up & Go Test

Considering cueing, GMWT and TUG showed somewhat higher ICC and/or lower $MDC_{95\%}$ when cueing was allowed or more extensive. In contrast, ICC were slightly higher and/or $MDC_{95\%}$ were lower for no cueing or less extensive cueing in FR, short distance WT, and short distance walking speed (see Table 18).

Table 18. Subgroup analysis of test-retest reliability considering cueing

No cueing	**Verbal cueing or verbal and visual/tactile cueing**	**More extensive cueing including physical assistance**
FR	ICC=0.84 $MDC_{95\%}$=15.4 % (Suttanon et al., 2011)	ICC=0.81 $MDC_{95\%}$=68.9 % (Muir-Hunter et al., 2015)

GMWT	ICC=0.57-0.96 $MDC_{95\%}$=31.2-225.7 % (Bossers, van der Woude et al., 2014)		ICC=0.96-0.99 $MDC_{95\%}$=19.6-33.3 % (H.-S. Lee et al., 2017)
TUG		ICC=0.76-0.96 $MDC_{95\%}$=23.3-39.6 % (Blankevoort et al., 2013; Suttanon et al., 2011; Thomas & Hageman, 1999)	ICC=0.72-0.99 $MDC_{95\%}$=15.8-30.0 % (H.-S. Lee et al., 2017; Muir-Hunter et al., 2015; Ries et al., 2009)
Short distance WT		ICC=0.80-0.95 $MDC_{95\%}$=31.6-41.5 % (Blankevoort et al., 2013; Thomas & Hageman, 1999)	ICC=0.85 $MDC_{95\%}$=84.3 % (H.-S. Lee et al., 2017)
Short distance walking speed	ICC=0.95-0.96 $MDC_{95\%}$=10.2-12.0 % (Wittwer et al., 2008)	ICC=0.50-0.95 $MDC_{95\%}$=31.6-48.3 % (Blankevoort et al., 2013; McGough et al., 2013; Suttanon et al., 2011)	ICC=0.85-0.98 $MDC_{95\%}$=25.5-84.3 % (H.-S. Lee et al., 2017; Ries et al., 2009)

FR: Functional Reach Test, GMWT: Groningen Meander Walking Test, ICC: intraclass correlation coefficient, $MDC_{95\%}$: percentage minimal detectable changes at 95 % confidence interval, TUG: Timed Up & Go Test, WT: walk tests

5.2.4.4 Frequency of use and effect sizes of motor assessments applied in previous randomised controlled trials

TUG, BBS, 5x STS, POMA, 30s CST, and instrumented gait analysis, were the most frequently applied assessments, utilised in six to sixteen RCT. We were only able to calculate effect sizes for twelve studies, as F/t statistics and/or standard deviations of baseline-post differences were infrequently reported. Effect sizes were large for FR, BBS, POMA, TUG, instrumented gait analysis, 5x STS, ACSID, and 30s CST (see Table 13/Additional file 9 for motor assessments identified during first search without available information on psychometric properties).

5.2.4.5 Summary and derivation of recommendations

Aiming to derive comprehensive recommendations on motor assessments for IWD, we combined the results of primary and secondary outcomes for each physical domain as summarised in Table 19.

Table 19. Summary of outcomes to derive recommendations for motor assessments for individuals with dementia

Motor assessment	Inter-rater reliability		Test-retest reliability		Frequency of use	Time*group interaction effect size
	relative	absolute	relative	absolute		
Balance						
FICSIT-4	?	?	0	-	-	?
Modified Clinical Test of Sensory Interaction of Balance	?	?	0	-	-	?
Limits of Stability	?	?	-	-	-	?
Physiomat-Trail-Making Task	?	?	0	?	-	?
Physiomat-Follow-The-Ball Task	?	?	0	?	-	?
FR	0	?	0	0	0	+
Hill Step Test	?	?	0	0	0	?
Step Quick Turn Test	?	?	-	-	-	?
Figure of Eight Test	?	?	0	-	-	?
GMWT	0	0	+	0	-	?
BBS	+	0	+	0	+	+
Modified BBS	?	?	?	?	-	?
POMA	?	?	?	?	+	+
Mobility and gait						
TUG	+	+	+	+	+	+
Cognitive TUG	?	?	-	0	0	?
Manual TUG	?	?	+	0	-	?
6m WT	0	0	0	-	0	?
4m WT	0	-	0	-	0	0
Instrumented gait analysis	?	?	0	0	+	+
Strength						
5x STS	?	?	0	0	+	+
STS on NeuroCom Balance Master	?	?	-	-	-	?
ACSID	+	?	?	?	-	+
30s CST	0	0	0	-	+	+
Handgrip dynamometer	?	?	+	-	0	-
Maximum isometric strength assessed with dynamometers	?	?	0	?	0	?

Endurance						
6min WT	0	?	+	0	0	?
Functional performance						
SPPB	?	?	?	?	0	0
E-ADL Test	?	?	0	?	0	?

4m WT: 4-metre walk test, 5x STS: Five Times Sit-to-Stand Test, 6m WT: 6-metre walk test, 6min WT: 6-minute walk test, 30s CST: 30-second chair stand test, ACSID: Assessment of Compensatory Sit-to-Stand Maneuvers in People With Dementia, BBS: Berg Balance Scale, E-ADL Test: Erlangen Test of Activities of Daily Living, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, FR: Functional Reach Test, GMWT: Groningen Meander Walking Test, POMA: Performance Oriented Mobility Assessment, SPPB: Short Physical Performance Battery, STS: Sit-to-Stand, TUG: Timed Up & Go Test

Relative reliability: -=insufficient, 0=sufficient, very low/low quality of evidence, +=sufficient, moderate/high quality of evidence, ?=not investigated

Absolute reliability: -=insufficient/unacceptable, 0=indeterminate/inconsistent/sufficient, very low/low quality of evidence, +=sufficient, moderate/high quality of evidence, ?=not investigated

Frequency of use: -=1 randomised controlled trial, 0=2-5 randomised controlled trials, +=> 5 randomised controlled trials

*Time*group interaction effect size*: -=no effect, 0=at least one trial with small or medium effect, +=at least one trial with large effect, ?=could not be calculated/not reported

Considering all information on primary and secondary outcomes, the derived recommendations include the following motor assessments:

- Balance: FR, GMWT (time), BBS, and POMA
- Mobility and gait: TUG and instrumented gait analysis to assess spatiotemporal gait parameters
- Strength: STS assessments with more than one repetition
- Endurance: 6min WT
- Functional Performance: No recommendation possible, due to insufficient research on psychometric properties

These recommendations are based on several outcomes rated in the highest category or one outcome rated in the highest and at least two in the second category (see Table 19).

5.2.5 Discussion

We addressed the purpose of this systematic review to quantitatively examine motor assessments for IWD by comprehensively analysing psychometric properties (primary outcome), frequency of use, and effect sizes (secondary outcomes) in a two-stage literature search. Recommendations on motor assessments are based on primary and

secondary outcomes. Additionally, we analysed the influence of severity and aetiology of dementia and cueing on test-retest reliability.

5.2.5.1 Findings on primary and secondary outcomes

The systematic search identified only few investigations examining validity, internal consistency, and intra-rater reliability of motor assessments in IWD. Thus, we were not able to draw further conclusions or consider these outcomes for deriving recommendations. Summarizing findings for inter-rater reliability shows sufficient relative inter-rater reliability and relatively low $MDC_{95\%}$ of considered motor assessments. Hence, they are objective measures to determine motor performance in IWD. Motor assessments analysing time in tasks of short duration, such as 4m WT, should, however, be treated with caution, as small measurement errors may significantly influence absolute inter-rater reliability. With respect to test-retest reliability, the majority of identified investigations observed sufficient relative test-retest reliability, while absolute test-retest reliability was mainly indeterminate or unacceptable. This supports their usage to investigate changes on a group level, but does not allow assessing intra-individual changes (Blankevoort et al., 2013; Bruton et al., 2000; Fox et al., 2016). Moreover, decreasing test-retest reliability from between-day, through within-day, to within-session investigations may be related to fluctuating daily forms in IWD. We expect that characteristics of daily form, such as mood or motivational aspects, remain relatively constant within short intervals, while they potentially alter with increasing time. More research is necessary to develop criteria to determine daily form, aiming to ensure comparable conditions in longitudinal investigations. Besides, fluctuating daily forms in IWD may have contributed to observed unacceptable absolute test-retest reliability. Other explanations refer to high intra-individual variability in IWD and related inappropriate or naive selection of metrics, which do not account for this variability.

Regarding frequency of use, previous trials predominately applied clinical motor assessments established in healthy older adults or various clinical populations, while those considering specific characteristics of IWD such as GMWT, Physiomat, or ACSID, were less frequently applied. This may be related to their first introduction between 2014 and 2018. Due to insufficient information in previous RCT, we were only able to determine time*group interaction effect sizes for 38 % of analysed motor assessments.

Based on large effect sizes reported in at least one RCT, we assumed sensitivity to change for most of these assessments.

5.2.5.2 Findings on influence of severity and aetiology of dementia and cueing on test-retest reliability

Considering severity of dementia, we expected decreasing test-retest reliability with increasing cognitive impairment. This assumption was true for FICSIT-4 and GMWT but not for all assessments. Severity of dementia may only influence specific assessments, for example those with complex instructions or assessing outcomes frequently impaired in IWD, such as balance (Allan et al., 2005). Unexpectedly, we observed increasing test-retest reliability with increasing severity of dementia for BBS, 6m WT, modified 30s CST, and 5x STS. However, these observations were only based on single studies, which partly differed in characteristics, such as aetiology of dementia.

Regarding the aetiology of dementia, test-retest reliability of BBS and up and go tasks was lower for AD than for various or not reported types. Both assessments consist of several short tasks and include multi-step instructions. Compared to other aetiologies, individuals with AD may have more difficulties in understanding and/or remembering such instructions, which potentially influences test-retest reliability (H.-S. Lee & Park, 2017; Muir-Hunter et al., 2015; Orange, Molloy, Lever, Darzins, & Ganesan, 1994). In contrast, test-retest reliability of walking speed was higher in AD which could be related to later occurring gait impairments in AD (Valkanova & Ebmeier, 2017). Additional research on aetiologies, however, is required to understand lower test-retest reliability of STS tasks and higher test-retest reliability of maximum isometric strength assessed with dynamometers in AD.

Analysing the influence of cueing on test-retest reliability revealed higher test-retest reliability when cueing was allowed or more extensive for GMWT and TUG, which are assessments consisting of unfamiliar or several short tasks. Cueing possibly stabilises motor performance by supporting impaired cognitive performance and thus improves test-retest reliability. In contrast, short distance WT, for which test-retest reliability was higher when cueing was not allowed or less extensive, are close to everyday life, include single-stage tasks, and consider well automated movement processes not re-

quiring additional cognitive support. Accordingly, cueing rather may distract IWD leading to destabilised performance decreasing test-retest reliability. No explanation for the same association in FR is available.

Based on these observed influences, we derived the following suggestions:

- Put emphasis on simple instructions, especially for IWD with advanced stages or AD.
- Consider individual cognitive and motor deficits, when selecting motor assessments.
- Only use cueing for motor assessments where it is inevitable.

5.2.5.3 Recommendations and need for future research

Recommendations for balance assessments include FR, GMWT (time), BBS, and POMA. Due to infrequent use and insufficient research on psychometric properties, feasibility and sensitivity to change of GMWT and psychometric properties of POMA, both assessments require further investigation. Focusing on mobility and gait, we suggest to apply TUG and spatiotemporal gait parameters assessed with instrumented gait analysis. Comparing different gait analysis systems, NeuroCom Balance Master, however, seems to be less suitable. Despite insufficient or equivocal results, future research should investigate short distance WT of different distances, as instrumented gait analysis systems may not be available for all studies. Considering strength, we suggest to apply STS assessments comprising more than one repetition, which, however, predominately determine functional performance of lower limbs. Thus, further evaluation of strength assessments including upper limb strength and measures allowing conclusion on actual strength performance are required. Moreover, we suggest to use the 6min WT as an endurance assessment for IWD. Future research on endurance assessment, however, is crucial since this was the only identified assessment. As information on psychometric properties is insufficient, we are not able to recommend any functional performance assessment. Based on secondary outcomes some indications are available for SPPB. However, psychometric properties of SPPB and other functional performance assessments need to be investigated in future studies.

5.2.5.4 Comparison with state of research

Recommendations of motor assessments in this review are largely in line with those of previous reviews (Bossers et al., 2012; McGough et al., 2019). Small discrepancies may be related to distinctions in identified assessments and studies, different prioritisation of considered outcomes, and divergent criteria for good measurement properties. Additionally, this review, consistently to Fox et al. (Fox et al., 2016), determined sufficient relative test-retest reliability for the majority of motor assessments in IWD, but remarked high $MDC_{95\%}$ reflecting unacceptable absolute test-retest reliability.

Similarly, motor assessments recommended in this review are mainly in line with those elaborated in a qualitative approach (Trautwein, Barisch-Fritz et al., 2019). However, FICSIT, 6m WT, SPPB, and Physical Performance Test were rated appropriate in the qualitative approach, but could not be recommended based on quantitative outcomes as they were infrequently used or insufficiently investigated. Further discrepancies on FR, which was rated inappropriate but can be recommended based on quantitative outcomes, require additional examination. Moreover, some general indications, related to consideration of specific characteristics and cueing are consistently suggested. Accordingly, this review largely sustains the recommendations elaborated in a qualitative approach.

5.2.5.5 General considerations on primary and secondary outcomes

The interpretation of findings regarding psychometric properties is challenging as there are no firm criteria for acceptable reliability in literature (Bruton et al., 2000). Regardless of concrete criteria, ICC do not only reflect relative reliability but also can be related to sample size or variability in the sample (Koo & Li, 2016). Accordingly, trial-to-trial consistency can be poor, despite high ICC. Thus, it is advised not to focus on single estimates of reliability and to additionally consider absolute reliability (Blankevoort et al., 2013; Bruton et al., 2000). Due to lack of information on minimal important change of motor assessments in IWD, we could scarcely apply COSMIN criteria for absolute reliability. Besides, Smidt et al. (Smidt et al., 2002) arbitrarily defined that a difference of 10 % in minimal detectable change would be acceptable. Other research groups referred to them and introduced another cut-off of 30 % without any justification (Huang et al., 2011; H.-S. Lee et al., 2017). In absence of other criteria, we adopted this cut-

off of 30 % to identify unacceptable $MDC_{95\%}$ but not to conclude on sufficient absolute reliability.

Frequency of use and effect sizes do not necessarily allow conclusions to be drawn on quality of motor assessments and should not be overestimated. Regardless of appropriateness and meaningfulness, researchers may decide to apply motor assessments as they are commonly used or easy to utilise. Nonetheless, frequency of use can provide indications about feasibility of motor assessments, which is based on the assumption that unfeasible motor assessments do not disseminate as good as feasible ones. Comparably, effect sizes can provide information on sensitivity to change, but are also dependent on effectiveness of interventions.

5.2.5.6 Strengths and limitations

To our knowledge, this is the first systematic review utilising a comprehensive approach combining different outcomes of previous reviews by performing an extensive two-stage literature search. We need to state potential risk of bias regarding the selection of considered motor assessments. Due to restricting the analysis of motor assessments to those applied in RCT, some assessments may be missing. Furthermore, large heterogeneity of included psychometric property studies limits the meaningfulness of derived recommendations. As psychometric properties are potentially influenced by various determinants, such as sample size, sample characteristics including severity and aetiology of dementia, cueing, test-retest interval, or considered outcomes, we cannot ensure that the deductions on psychometric properties are true and not randomly caused by differing determinants. Therefore, false assumptions, undetected influences or relations, and random observations may have occurred. Similarly, the consideration of several influences on test-retest reliability only allows rough estimations, which could be also affected by heterogeneity of analysed studies. Moreover, insufficient information on execution of motor assessments, severity and aetiology of dementia, and cueing in available investigations impeded detailed analyses and limited meaningfulness of observations. Accordingly, the elaborated recommendations should be used with care and further research investigating psychometric properties and dementia specific influences on test-retest reliability is required.

5.2.6 Conclusion

Despite the necessity for further research in various areas, this review establishes an important foundation for future investigations. Additionally, direct implications for studies determining effectiveness of physical activity on motor performance in IWD can be derived. However, elaborated recommendations cannot be considered as final conclusions since the analysis of primary and secondary outcomes reveals several challenges and areas of insufficient research, and only focus on quantitative aspects. Furthermore, new assessments, especially developed for IWD, are required. Such assessments can be based on prior tasks but should consider specific characteristics of IWD. Additionally, it is of high importance to standardise motor assessments and cueing to ensure comparability between studies. Herein, standardisation refers to selection and performance procedures of motor assessments and external cues. Currently, a wide range of motor assessments (e.g. previous RCT applied 19 different balance assessments) with different performance procedures (e.g. different ratings or modifications) as well as various external cues (e.g. clearly defined verbal cues vs. as much assistance as needed) are frequently applied to determine the same motor functions or quantities. Accordingly, recommendations on specific motor assessments as well as indications on assessment procedures elaborated in quantitative and qualitative (see Trautwein, Barisch-Fritz et al., 2019) approaches are important to improve standardisation. Evidence on effectiveness of physical activity can contribute to gain access to physical activity interventions and thereby positively influence quality of life in IWD. Determining evidence, however, is not possible without appropriate, sensitive, valid, reliable, and standardised motor assessments, which consider the individual characteristics of single individuals.

Abbreviations

4m WT: 4-metre walk test; 5x STS: Five Times Sit-to-Stand Test; 6m WT: 6-metre walk test; 6min WT: 6-minute walk test; 30s CST: 30-second chair stand test; ACSID: Assessment of Compensatory Sit-to-Stand Maneuvers in People With Dementia; AD: Alzheimer's disease; BBS: Berg Balance Scale; COSMIN: COnsensus-based Standards for the selection of health Measurement INstruments; E-ADL Test: Erlangen Test of Activities of Daily Living; FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4; FR: Functional Reach Test; GMWT: Groningen Meander Walking Test; ICC: intraclass correlation coefficient/s; IWD: individuals with dementia; $MDC_{95\%}$: percentage minimal

detectable change/s at 95 % confidence interval; MMSE: Mini-Mental State Examination; POMA: Performance Oriented Mobility Assessment; RCT: randomised controlled trial/s; SPPB: Short Physical Performance Battery; STS: Sit-to-Stand/sit-to-stand; TUG: Timed Up & Go Test; WT: walk test/s

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Funding

This project is financially supported by the Dietmar Hopp Stiftung (St. Leon-Rot, Germany). The sponsor does not have any role in the design of the study, neither in its execution, the collection, analysis or interpretation of data, the decision to submit results nor in writing the report.

Authors' contributions

ST, BB, PM, and AH performed the systematic researches. BB and AW were consulted as third reviewer, if no consensus could be reached. ST, PM, and AH performed data extraction. ST and PM analysed results. All authors were involved in planning and preparing the systematic review. ST, PM, and BB wrote the manuscript. All authors provided critical feedback and approved the final manuscript.

Acknowledgements

We would like to thank Emily Cooke for her linguistic assistance on behalf of the authors. We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Karlsruhe Institute of Technology.

Additional files

Additional file 1: Search term first search

Additional file 2: Formulas for calculating time*group interaction effect sizes

Additional file 3: Search term main search

Additional file 4: Formulas for calculating minimal detectable change at 95 % confidence interval

Additional file 5: Study characteristics first search

Additional file 6: Study characteristics main search

Additional file 7: Data extraction first search

Additional file 8: Data extraction main search

Additional file 9: Description, frequency of use, and effect sizes of motor assessments applied in previous randomised controlled trials without available information on psychometric properties

References

All references of manuscript II are included in the List of References at the end of this thesis.

5.3 Study design to investigate the effectiveness of physical activity on motor performance in individuals with dementia

Manuscript III

Summary: Due to several methodological limitations and risk of bias, evidence on the effectiveness of physical activity on motor and gait performance in IWD cannot be ensured (Blankevoort et al., 2010; Hernández et al., 2015; Lam, Huang et al., 2018). Accordingly, recent reviews indicate the need for further high-quality studies (Blankevoort et al., 2010; Lam, Huang et al., 2018). Considering the criteria for high-quality study designs derived from limitations criticized in these recent reviews (see Table 8), manuscript III aims to establish a high-quality methodological approach to investigate the effectiveness of physical activity on motor and gait performance in IWD. Herein, it provides an example including valuable indications to answering research question A3.

Published in: JMIR Research Protocols

Published on: March 3rd, 2017

Reference:

Trautwein, S., Scharpf, A., Barisch-Fritz, B., Niermann, C., & Woll, A. (2017). Effectiveness of a 16-week multimodal exercise program on individuals with dementia: Study protocol for a multicenter randomized controlled trial. *JMIR Research Protocols*, *6*(3), e35. https://doi.org/10.2196/resprot.6792

Effectiveness of a 16-Week Multimodal Exercise Program on Individuals With Dementia: Study Protocol for a Multicenter Randomized Controlled Trial[10]

Sandra Trautwein, MA; Andrea Scharpf, MA; Bettina Barisch-Fritz, Dr rer nat; Christina Niermann, Dr Phil; Alexander Woll, Prof Dr

Institute of Sports and Sports Science, Karlsruhe Institute of Technology, Karlsruhe, Germany

Corresponding Author: Sandra Trautwein

5.3.1 Abstract

Background: The increasing prevalence of dementia in the next decades is accompanied by various societal and economic problems. Previous studies have suggested that physical activity positively affects motor and cognitive skills in individuals with dementia (IWD). However, there is insufficient evidence probably related to several methodological limitations. Moreover, to date adequate physical activity interventions specifically developed for IWD are lacking.

Objective: This study aims to investigate the effectiveness of a multimodal exercise program (MEP) on motor and cognitive skills in IWD in a high-quality multicenter trial.

Methods: A multicenter randomized controlled trial with baseline and postassessments will be performed. It is planned to enroll 405 participants with dementia of mild to moderate stage, aged 65 years and older. The intervention group will participate in a 16-week ritualized MEP especially developed for IWD. The effectiveness of the MEP on the primary outcomes balance, mobility, and gait will be examined using a comprehensive test battery. Secondary outcomes are strength and function of lower limbs, activities of daily living, and cognition (overall cognition, language, processing speed, learning and memory, and visual spatial cognition).

Results: Enrollment for the study started in May 2015. It is planned to complete postassessments by the beginning of 2017. Results are expected to be available in the first half of 2017.

Conclusions: This study will contribute to enhancing evidence for the effects of physical activity on motor and cognitive skills in IWD. Compared to previous studies, this study is characterized by a dementia-specific intervention based on scientific knowledge, a

[10] Some minor formal adaptions were made to the version of manuscript III presented in this thesis to ensure uniform formatting.

combination of motor and cognitive tasks in the intervention, and high standards regarding methodology. Findings are highly relevant to influence the multiple motor and cognitive impairments of IWD who are often participating in limited physical activity.

Trial Registration: German Clinical Trials Register DRKS00010538; https://drks-neu.uniklinik-freiburg.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00010538 (Archived by WebCite at http://www.webcitation.org/6oVGMbbMD)

Keywords: physical activity; dementia; postural balance; gait; activities of daily living; cognition; exercise

5.3.2 Introduction

Dementia is one of the most frequently occurring diseases in the elderly (Berr, Wancata, & Ritchie, 2005), and the World Health Organization has declared dementia a public health priority (World Health Organization, 2012). The current prevalence of dementia is estimated at 47 million worldwide (Prince, Guerchet, & Prina, 2015) and will presumably increase because of expected demographic changes (Sosa-Ortiz, Acosta-Castillo, & Prince, 2012). This increasing prevalence (expected 135 million in 2050; Prince, Guerchet, Prina, & Alzheimer's Disease International, 2013) will be accompanied by several societal and economic problems including rising disease-related costs and increasing demands for caregiving (World Health Organization, 2012).

Dementia is a syndrome which comprises several different types of usually chronic and progressive diseases of the brain (eg, Alzheimer disease or vascular dementia; World Health Organization, 1992). It encompasses diverse impairments and symptoms which affect individuals with dementia (IWD) in different ways depending on dementia type (Chew-Graham & Ray, 2016). According to the International Statistical Classification of Diseases and Related Health Problems, Tenth Edition (ICD-10) (World Health Organization, 1992), a diagnosis of dementia minimally requires the following symptoms: an impaired memory, further cognitive disturbances, and noncognitive disorders such as disturbed emotional control. These impairments potentially influence activities of daily living (ADL) (World Health Organization, 1992) accompanied by an increasing loss of independence to a greater or lesser extent (Martyr & Clare, 2012). In addition, IWD suffer from motor and functional impairments such as affected gait and balance performance as well as transfer movements, which are not only reported in advanced stages (Allan et al., 2005; Manckoundia et al., 2006; van Iersel et al., 2004).

To date, there is no cure for dementia, and commonly used medications for treating the symptoms of dementia have side effects emphasizing the urgent need for nonpharmacological interventions (Groot et al., 2016). For instance, there is evidence that physical activity positively affects motor and cognitive skills of cognitively healthy elderly people (Angevaren, Aufdemkampe, Verhaar, Aleman, & Vanhees, 2008). Moreover, the number of studies analyzing this issue in IWD has increased (Christofoletti et al., 2008; Hauer et al., 2012; Kemoun et al., 2010; Rolland et al., 2007; Venturelli et al., 2011). For this sample, there are also systematic reviews and meta-analyses examining the effects of physical activity on balance, mobility, and gait as well as strength and ADL. Regarding balance, 3 of 5 reviews reported no or no clear benefit of physical activity (Brett et al., 2016; Littbrand et al., 2011; Potter et al., 2011) with largely varying effect sizes from small negative to large positive values (Blankevoort et al., 2010; Suttanon et al., 2010). Even if a positive effect of physical activity on mobility can be reported (Brett et al., 2016; Pitkälä, Savikko et al., 2013; Potter et al., 2011), the overall conclusion is inconsistent (Littbrand et al., 2011) with effect sizes ranging from small negative to large positive values (Blankevoort et al., 2010; Suttanon et al., 2010). Only a few reviews have considered specific aspects of gait function. One review has shown no to medium effect sizes for normal gait speed (Blankevoort et al., 2010). Reviews focusing on strength of lower limbs and ADL mainly reported improvements (Blankevoort et al., 2010; Brett et al., 2016; Forbes et al., 2015; Heyn, Johnson, & Kramer, 2008; Littbrand et al., 2011; Pitkälä, Savikko et al., 2013, 2013; Potter et al., 2011; Suttanon et al., 2010). However, the small number of high-quality studies and the large heterogeneity in methods used in these studies represent insufficient evidence regarding the effects of physical activity (Blankevoort et al., 2010; Forbes et al., 2015).

Reviews and meta-analyses examining the effects of physical activity on cognitive skills in IWD mainly assess overall cognition. Of 6 reviews and meta-analyses, 3 found no evidence for the benefit of physical activity on cognition in IWD (Forbes et al., 2015; Littbrand et al., 2011; Öhman et al., 2014) while the others found a positive overall effect (Brett et al., 2016, 2016; Farina et al., 2014; Groot et al., 2016). Groot et al. (2016) stated that overall effects of physical activity on cognition are comparable to the

effect size observed in meta-analyses examining the effectiveness of pharmacotherapy in IWD (Di Santo, Prinelli, Adorni, Caltagirone, & Musicco, 2013; Matsunaga, Kishi, & Iwata, 2015a, 2015b).

Most of the systematic reviews and meta-analyses suggest even if evidence is lacking that physical activity positively affects IWD, for example, in balance, mobility, and cognition. Their conclusions are that there is an urgent need for high-quality intervention studies (Blankevoort et al., 2010; Groot et al., 2016; Littbrand et al., 2011; Öhman et al., 2014; Suttanon et al., 2010). In their opinion, methodological shortcomings including insufficient reporting of methods and results and small samples as well as the use of inadequate outcome measures (Blankevoort et al., 2010; Farina et al., 2014; Groot et al., 2016; Öhman et al., 2014) could be responsible for the lack of conclusive evidence. Furthermore, Hauer et al. (2006) discussed that low effectiveness of existing physical activity interventions may explain negative or inconsistent findings in previous studies. It can be speculated that the effectiveness of existing training interventions is limited by inappropriate intensity, duration, type of training, lack of specific interventions, or individualization of training (Hauer et al., 2006).

This study will investigate the effects of a physical activity intervention on motor and cognitive skills. The intervention focuses on dementia-specific motor deficits and aims to influence the underlying motor performance, which depends on complex cognitive processes like integrating sensory information, central processing, or efferent motor output (Hüger et al., 2009). This reflects the close connection between cognitive and motor functions and could provide insights in disease progression (Alexander & Hausdorff, 2008). It is highly relevant for IWD to counteract and possibly reduce dementia-related motor deficits which typically result in distinct constraints of mobility-dependent quality of life as well as loss of independence and higher risk for falls (American Geriatrics Society, British Geriatrics Society, & American Academy of Orthopaedic Surgeons Panel on Falls Prevention, 2001; Guralnik, Ferrucci, Simonsick, Salive, & Wallace, 1995; Yümin, Şimşek, Sertel, Öztürk, & Yümin, 2011). Hence, primary outcomes are based on 3 considerations: dementia-specific motor deficits, relevance for everyday life, and measurement quality (direct and feasible measurements). Balance, gait, and mobility fulfill all requirements and influence quality of life (Allan et al., 2005; Telenius, Engedal, & Bergland, 2013; Vermeulen, Neyens, van Rossum, Spreeuwenberg, & Witte, 2011; Yümin et al., 2011). ADL are defined as secondary outcomes

because they are considered an entire construct related to several motor and cognitive skills. Thus, measuring ADL is more difficult and less objective than measuring balance, mobility, and gait. Further, we chose strength and function of lower limbs and cognition as secondary outcomes because of their expected influence on primary outcomes.

Aiming to overcome the above mentioned methodological limitations, we will realize a high-quality multicenter trial with a sustainable intervention close to everyday life. The following aims will be addressed.

Primary aim: to determine the effect of a multimodal exercise program (MEP) compared to conventional treatment (eg, medication, care, therapeutic applications) on balance, mobility, and gait. We hypothesize that a 16-week MEP in addition to conventional treatment affects balance, mobility, and gait in IWD more than the conventional treatment. Additionally, we will compare different subgroups (eg, according to sex, stage of dementia, or attendance).

Secondary aim: to investigate the influence of mediator and moderator variables on primary outcome measures. We assume that the effects of physical activity on balance, mobility, and gait are caused or influenced by changes in underlying motor and cognitive skills.

Comparably, we will investigate the effect of MEP on the secondary outcomes strength and function of lower limbs, ADL, and cognition as well as the effect of mediator and moderator variables on ADL. By addressing these aims, this study contributes to enhancing evidence concerning the effects of physical activity on motor and cognitive skills in IWD.

5.3.3 Methods

5.3.3.1 Study Design

The study design has been primarily defined to address the primary aim of the study on the effectiveness of a 16-week MEP. For this reason, we will perform a multicenter randomized controlled trial with baseline and postassessments and an allocation ratio of 2:1 for intervention (IG) and control group (CG), respectively. Ethical approval has been obtained from the ethics commission of the Karlsruhe Institute of Technology.

The study is retrospectively registered in the German National Register of Clinical Trials [DRKS00010538]. This study protocol considers guidelines and recommendations of the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) (Chan, Tetzlaff, Altman et al., 2013; Chan, Tetzlaff, Gøtzsche et al., 2013) and Consolidated Standards of Reporting Trials (CONSORT) statements (Boutron, Moher, Altman, Schulz, & Ravaud, 2008; Moher et al., 2010; Schulz, Altman, & Moher, 2010).

5.3.3.2 Participants

Participants for this study will be recruited in public, private, and charitable care facilities in southwestern Germany, in particular in the metropolitan region Rhein-Neckar and the district around Karlsruhe. All randomly selected care facilities offer inpatient care for approximately 60 to 300 residents and provide a common room where the intervention will be performed. A total of 3 recruitment periods with consecutive sampling within each care facility are planned.

Employees of care facilities will identify possible participants with the purpose to fulfill selection criteria.

Inclusion criteria include (1) diagnosis of dementia or suspected dementia (based on the assessment of the objective ICD-10 criteria by employees and the examination of cognitive abilities with Mini-Mental State Examination [MMSE]; Folstein et al., 1975), (2) Alzheimer disease, vascular dementia, or other primary dementia (all types caused by neurodegenerative or vascular diseases: eg, lewy body dementia or frontotemporal dementia; Reith & Muhl-Benninghaus, 2015), (3) mild to moderate stage of dementia (MMSE 10-24), (4) age above 65 years, (5) walking ability of approximately 10 meters with or without walking aid, and (6) clearance by general practitioner.

Exclusion criteria include (1) secondary dementia (all types resulting from organic illness or injury: eg, toxic substances or brain injuries; Reith & Muhl-Benninghaus, 2015), (2) other severe cognitive impairments, (3) other severe neurological disease, (4) other severely acute diseases, and (5) severe motor impairments.

Potential participants will receive a comprehensive information letter and an informed consent form, which will be signed by individuals or their legal guardians prior to the study. The informed consent along with clearance of participant’s general practitioner

allow scheduling of baseline assessments where eligibility will be verified according to the inclusion and exclusion criteria. Flow of participants is illustrated in Figure 5.

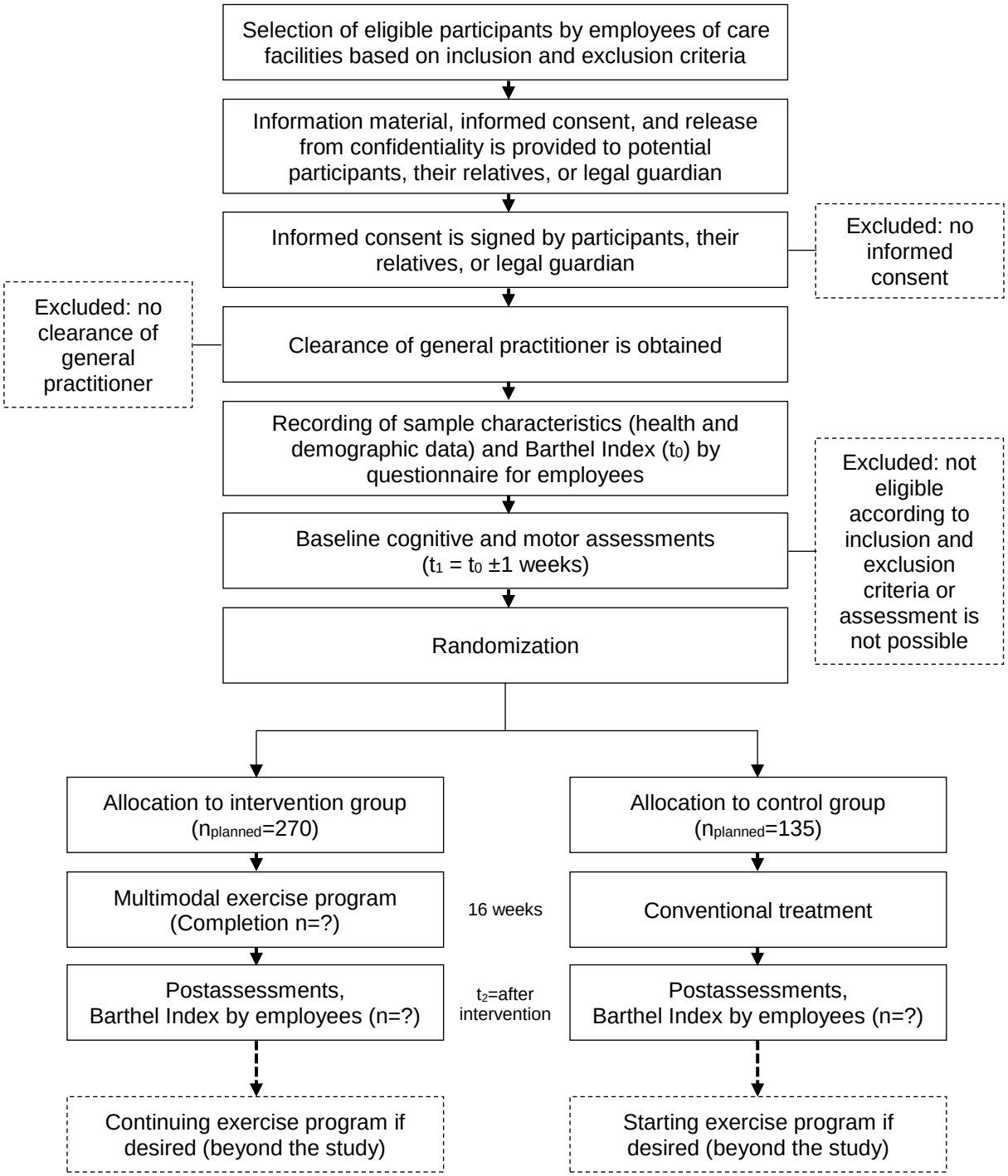


Figure 5. Flow of participants.

5.3.3.3 Intervention

The intervention is specifically developed for this study based on theoretical considerations, results of a pilot study (Thurm et al., 2011), and a literature review (Scharpf et al., 2013). The combination of motor and cognitive tasks used in the MEP aims to enhance the effectiveness of physical activity on cognition. This is theoretically supported by findings in healthy older adults showing that the combination of both yields larger effects on cognition than using each alone (Fabre et al., 2002). The pilot study (n=19) aimed to prove feasibility of the intervention and allowed first insights regarding the effectiveness. After a 10-week intervention, IG showed no significant changes in Alzheimer Disease Assessment Scale–Cognitive Subscale (ADAS-Cog, German version; Ihl & Weyer, 1993) sum score but significant improvements in subscore orientation/praxis. In contrast, we found a significant decline in ADAS-Cog sum score of CG (Thurm et al., 2011). Moreover, IG showed significant improvements in get-up-and-go test whereas CG did not significantly improve (unpublished results). The literature review aimed at giving recommendations for designing interventions for IWD. Analyzed studies showed that a physical activity intervention for IWD should at least last 4 months with 2 to 3 sessions of 45 to 60 minutes per week. Moreover, interventions focusing on several motor skills (eg, endurance, strength, balance) seemed to be more effective than interventions with only 1 task (Scharpf et al., 2013). Hence, the 10-week intervention of the pilot study has been revised for the current study. The revision, which aimed to provide a balanced MEP with specific, adequate, and intensity-demanding tasks, comprises adjustments of contents (motor qualities as well as connection between motor and cognitive tasks) and intervention duration (extension to 16 weeks).

The MEP will be guided by 2 skilled instructors with experience in sports science and performed as group training mainly in a seated position. A group will consist of a maximum of 12 participants and will be joined by familiar caregivers to support the instructors if needed. The underlying didactic concept focuses on specific needs and characteristics of IWD and includes increased supervision realized by 2 instructors, adaptation to the cognitive level of participants, adjusted communication (eg, simple language, nonverbal aspects), ritualization to give orientation and familiarity, and adequate complexity by simple and well-structured cognitive and motor tasks.

To ensure high standards and comparability, each session is planned in detail and all instructors participate in a special training focusing on structure and contents of MEP as well as special demands resulting from the characteristics of IWD. A detailed training manual is provided for instructors, and the adherence to this manual will be emphasized. To ensure standardization, all tasks are described precisely and photographs are provided.

Providing a sense of security is an important aspect realized by ritualization. To satisfy this ritualization, the general sequence is identical for all sessions including an imagination of experienced journeys. Each session is divided into 3 parts: arrival, destination, and departure. Whereas arrival and departure remain consistent over the whole intervention period, a new travel destination is selected every time. A total sample session of MEP is found in Multimedia Appendix 1.

The arrival as beginning ritual of each training session takes about 5 to 7 minutes and aims to prepare participants for the following main part. Tasks for mobilization and stimulation of the cardiovascular system are linked to cognitive activation.

The main part of MEP is the destination (about 35 minutes) which includes tasks for strength (43 %), balance (25 %), endurance (16 %), flexibility (13 %), and not further specified tasks (3 %; see Figure 6). In addition, cognitive tasks are incorporated to stimulate memory, attention, language, and executive functions. Tasks are carried out with medium to submaximal intensity. Throughout the intervention, there will be a progression concerning intensity as well as motor and cognitive requirements. Examples of different motor and cognitive tasks as well as examples for their progression are given in Table 20.

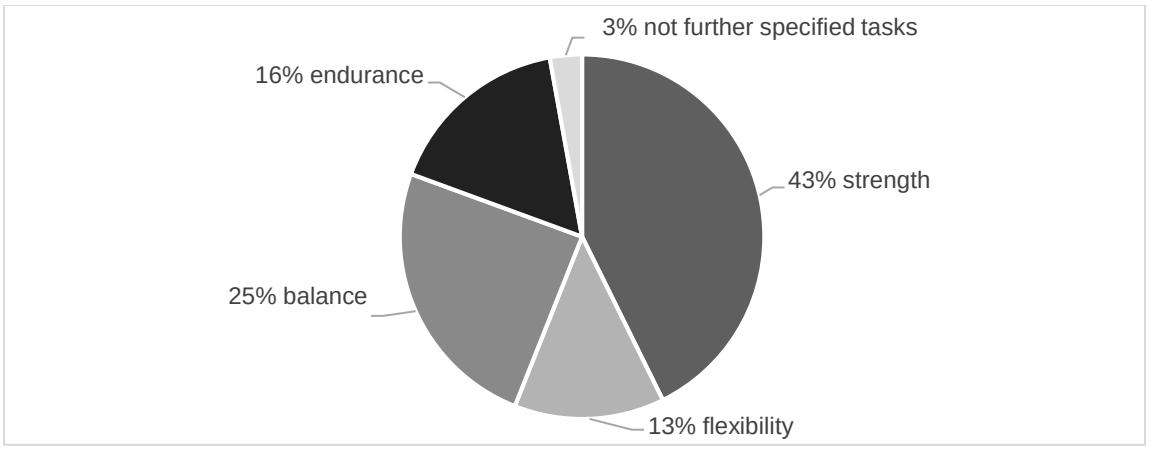


Figure 6. Distribution of motor qualities within the main parts of the multimodal exercise program.

Table 20. Examples of motor and cognitive tasks of the multimodal exercise program and their progression

		Simple performance	**Progressive performance developed within the 16 weeks**
Strength	Imagination/journey	Mediterranean cruise – aquafitness on the deck of the ship	Circus – task after tightrope dance
	Starting position	Seated, arms stretched above head	Standing upright behind chair, arms stretched above head
	Motor task	Lateral flexion with pool noodle	Lateral flexion with rope
	Sets and repetitions	3 sets with 2 repetitions for each side	2 sets with 3 repetitions for each side
	Muscle activity	Upper limbs and core	Upper limbs, core, and lower limbs
	Cognitive task	No additional cognitive task	Answering questions about circus performances (eg, Have you ever been to a circus? If yes: Which was the best circus act? If no: What do you think would be the most interesting thing if you visited a circus?)
Balance	Imagination/journey	Safari in Namibia – washing an elephant	World trip – washing an elephant
	Starting position	Seated, 1 arm is horizontally stretched, flexion in hip joint to shift body weight forward	Standing upright behind chair, one arm is horizontally stretched, flexion in hip joint to shift body weight forward
	Motor task	Slow and large arm movements in horizontal plane holding a small sandbag while leaning to left and right sides	Slow and large arm movements in horizontal plane holding a small sandbag while leaning to left and right sides
	Cognitive task	Answering questions about elephants (eg, Have you ever seen an elephant? Are there different kinds of elephants? What are the differences?)	Counting to 180 in steps of 6 (change hands at 90)
	Duration/repetitions	1 minute/approximately 10 repetitions per side	Approximately 1:30 minutes/15 repetitions per side
Endurance	Imagination/journey	Soccer World Cup – walking to the soccer training	On a treasure island – walking downhill through the jungle
	Starting position	Seated	Standing upright behind chair
	Motor task	“Walking” in seated position – lifting legs with active use of arms	“Walking” on the spot – lifting legs with active use of arms (if possible)
	Duration	1 minute	3 minutes
	Cognitive task	Answering questions about soccer and its rules (eg, Who knows some soccer rules? Do you know how many referees there are during a game?)	Naming animals living in the jungle. If a participant repeats an animal he or she is asked to name another one

Flexibility	Imagination/journey	Safari in Namibia – wood chopping for a campfire	Olympic Games – laola wave of the audience
	Starting position	Seated	Standing upright behind chair
	Motor task	Extension and flexion of the trunk, bringing arms in extension with maximal personal range of motion	Extension and flexion of the trunk, bringing arms in extension with maximal personal range of motion (try to increase range of motion)
	Set and repetitions/ duration	3 sets with 10 repetitions (5 repetitions slow, 5 repetitions fast)	No repetitions defined, duration 3 minutes
	Cognitive task (example)	Performing in the same rhythm synchronous with other participants, 5 slow hits, 5 faster hits	Learning 3 different signals: 1= moving fast, 2= moving slow, 3= change direction of laola wave, performing according to signals

The departure takes about 5 minutes and aims to cool down and relax the body while leading participants out of imagination and back into reality. Similarly to the arrival, instructors guide participants through fixed sequences.

The MEP takes place twice a week on nonconsecutive days over a period of 16 weeks. Each session lasts 60 minutes with motor and cognitive tasks taking about 45 minutes to ensure sufficient time for rests and explanations. Prior to the first session, a social gathering session is held aiming for an initial familiarization and information acquisition with regard to participants and care facilities. Attendance and adherence of participants will be documented by instructors for each session. Adherence will be assessed using a short formula to rate attention, participation, motivation, and behavior of each participant.

Conventional treatment comprising, for instance, medication, care, or therapeutic applications is individually tailored and will be continued in all included participants of CG as well as IG.

5.3.3.4 Outcomes

5.3.3.4.1 Determination of Outcomes

Primary outcomes refer to the motor qualities balance, mobility, and gait. Secondary outcomes are other motor variables such as strength and function of lower limbs and ADL as well as cognitive variables assessing overall cognition, language, processing speed, learning and memory, and visual spatial cognition. All outcome parameters are listed in Table 21. The aim of this study is to investigate changes in outcomes between

IG and CG. Furthermore, the focus is on differences in all outcome variables between baseline and 16-week postassessment.

Table 21. Primary and secondary outcome parameters

	Outcome	Assessments (at baseline and 16-week postassessment)
Primary outcomes	Balance	Frailty and Injuries: Cooperative Studies of Intervention Techniques 4 (FICSIT-4) (Rossiter-Fornoff et al., 1995)
	Mobility	Timed Up and Go test (Podsiadlo & Richardson, 1991) 6-meter walk test (Graham, Ostir, Kuo, Fisher, & Ottenbacher, 2008)
	Gait	Gait analysis using GAITRite: temporal and spatial gait parameters (gait speed, cadence, cycle time, step length, step width, gait variability, single support, and double support) - Walking with normal speed - Walking with normal speed and the task counting backwards from 50 - Walking with normal speed and the task naming animals
Secondary outcomes	Lower limb strength	Modified 30-second chair-stand test (Blankevoort et al., 2013; Jones et al., 1999)
	Lower limb function	Short physical performance battery (Guralnik, Simonsick et al., 1994)
	Activities of daily living	Barthel Index (German version according to Hamburger Einstufungsmanual; Lübke, Meinck, & Renteln-Kruse, 2004; Mahoney & Barthel, 1965) Erlangen Test of Activities of Daily Living (E-ADL-Test) (Graessel et al., 2009) 7-item physical performance test (Reuben & Siu, 1990)
	Overall cognition	Mini-Mental State Examination (MMSE) (Folstein et al., 1975)
	Language	Verbal fluency “category animals” Phonemic fluency “S-words”
	Processing speed	Trail Making Test A (Reitan, 1958, 1992)
	Learning and memory	California Verbal Learning Test, short version 1 (Elwood, 1995) Digit span forward and backward (Wilde, Strauss, & Tulsky, 2004)
	Visual spatial cognition	Clock drawing test (Shulman, Shedletsky, & Silver, 1986)

The primary and secondary outcomes have been discussed in an international expert panel consisting of 14 scientists from 7 institutions in 3 countries (Germany, Australia, and Netherlands) with the disciplines sports science (especially focusing on locomotion research, sports therapy, kinesiology, biomechanics, training science, physical education and health, diagnostics, evaluation, and sports psychology), geriatrics/gerontology, psychology, and physiology. Among these experts, a standardized testing procedure has been determined focusing on relevance of outcomes as well as validity, reliability, objectivity, and feasibility of recording methods. The selected outcomes and recording methods are common in geriatric assessments and have been frequently

used in previous studies examining IWD. However, it must be pointed out that most of recording methods regarding the motor qualities have not been developed for IWD. Feasibility of the test battery and recording procedure was tested in a sample of 20 participants prior to the current study. This pilot study proved feasibility of planned assessments in IWD.

Trained investigators with experience in sports science guide the baseline and postassessments in the care facilities. Prior to assessments, investigators participate in a special course to get detailed information about testing procedure and measurements. To standardize testing procedure and ensure comparability, a detailed testing manual is provided to which investigators are urged to strictly adhere. Accordingly, a detailed description of performing each assessment is given in Multimedia Appendix 2. Moreover, investigators will be educated about specific aspects of working with IWD.

5.3.3.4.2 Primary Outcomes

Static balance will be determined using the Frailty and Injuries: Cooperative Studies of Intervention Techniques 4 scale (FICSIT-4) (Rossiter-Fornoff et al., 1995). Mobility will be assessed using the timed Up and Go test (Podsiadlo & Richardson, 1991) and 6-meter walk test (Graham et al., 2008). The 6-meter walk test aims to capture normal gait speed. To reduce bias caused by the testing situation, participants are not explicitly informed about time keeping.

Temporal and spatial gait parameters will be analyzed using the electronic gait analysis system GAITRite (CIR Systems Inc, Franklin, NJ) with an active length of 4.88 meters, a spatial resolution of 1.27 centimeters, and a scan rate of 120 hertz. The following parameters are of special interest: gait speed, cadence, cycle time, step length, step width, gait variability, single support, and double support (as percentage of cycle time). Gait parameters are recorded for 3 different conditions: walking with normal speed, walking with normal speed and the task of counting backwards from 50, and walking with normal speed and the task of naming animals.

Changes in gait parameters caused by dual task will be calculated using the equation seen in Figure 7. The generated value represents dual-task costs indicating the better performance under dual-task condition the lower this value is (Abernethy, 1988; Schwenk et al., 2010).

$$\frac{dual\ task - single\ task}{single\ task} \cdot 100$$

Figure 7. Calculation of changes in gait parameters caused by dual task.[11]

5.3.3.4.3 Secondary Outcomes

Strength of lower limbs will be determined by modified 30-second chair-stand test. In this modified version participants are allowed to use their arms (Blankevoort et al., 2013; Jones et al., 1999), and the time to perform 5 repetitions is additionally measured. After a rest, fit participants complete a second trial without using arms with the same recording procedure as for the modified 30-second chair-stand test (including time for 5 repetitions). Function of lower limbs will be evaluated using the short physical performance battery, consisting of standing balance (Romberg, semitandem, tandem), gait speed, and 5 times sit-to-stand without using arms (Guralnik, Simonsick et al., 1994).

ADL will be determined using the Barthel Index (German version according to Hamburger Einstufungsmanual; Lübke et al., 2004; Mahoney & Barthel, 1965), Erlangen Test of Activities of Daily Living (E-ADL-Test) (Graessel et al., 2009), and 7-item physical performance test (Reuben & Siu, 1990). The Barthel Index will be completed by employees of the care facilities. To ensure standardized answers, employees receive a manual with detailed information. The E-ADL-Test and the 7-item physical performance test aim to practically examine ADL. Although the revalidation of the E-ADL-Test (Graessel et al., 2009; Luttenberger et al., 2012) showed that the tasks are too easy for mild dementia, for our target sample this test is considered as appropriate substantiated by the development for IWD. Furthermore, the E-ADL-Test is regarded as a valid and reliable instrument for assessing ADL of individuals with moderate to severe dementia (Graessel et al., 2009; Luttenberger et al., 2012).

Cognitive outcomes will be assessed using some subtests of the neuropsychological test battery Consortium to Establish a Registry for Alzheimer's Disease–Plus (CERAD-Plus) (J. C. Morris, Mohs, Rogers, Fillenbaum, & Heyman, 1988). Overall cognition will be determined using MMSE (Folstein et al., 1975). Language will be examined regarding verbal fluency “category animals” and phonemic fluency “S-words.” The first fluency

[11] Corrected.

task provides information about verbal rate and fluency, semantic memory, language, executive function, and cognitive flexibility (Lezak, 2012; J. C. Morris et al., 1989). The second task examines fluency in a more strategic manner rather than the semantic memory. Processing speed and visual scanning will be determined using the Trail Making Test A (Reitan, 1958, 1992). In addition to CERAD-Plus, the California Verbal Learning Test, short version 1 (except forced choice recognition; Elwood, 1995), and digit span forward and backward (Wilde et al., 2004) will be performed to assess learning and memory. Visual spatial cognition will be assessed using the clock drawing test (Shulman et al., 1986).

Moreover, body mass and height will be measured using a Seca 813 Robusta scale and Seca 213 stadiometer (Seca, Hamburg, Germany) with an accuracy of 0.1 kilogram and 0.1 centimeter, respectively.

5.3.3.4.4 Sample Characteristics

Further possible influencing variables including age, medication, or other diseases are recorded chronologically close to baseline assessments. Employees of the care facilities will be asked to complete the health and demographic data questionnaire and the Cumulative Illness Rating Scale (Linn, Linn, & Gurel, 1968) for each participant. The questionnaire includes sex, year of birth, diagnosis of dementia, severity of dementia, type of dementia, date of diagnosis, depression, severity of depression, number of medications, medications for dementia, antidepressants, and walking aids. A written consent to collect these data by employees of the care facilities will be obtained from participants or their legal guardian.

5.3.3.5 Sample Size

The required sample size was calculated via G*Power version 3.1.9.2 (Heinrich Heine University of Dusseldorf; Faul, Erdfelder, Lang, & Buchner, 2007), taking into account the following parameters: analysis of variance (ANOVA) for repeated measures, within-between interaction, small effect size (η^2=0.01, d=0.2; Jacob Cohen, 1988), 2-sided α-error of .05, power of .80 (1-β), and 2 groups and 2 measurements. The small effect size used for the calculation of required sample size is based on literature review and assumptions of relevant changes for IWD. Previous studies have reported high variation in the effect sizes of the primary outcomes balance, mobility, and gait. In their

review, Blankevoort et al. (2010) reported small negative to large positive effect sizes for balance (d=-0.24 to d=3.59) and functional mobility (d=-0.25 to d=2.37) as well as no to medium effect sizes for normal gait speed (d=-0.11 to d=0.50). These reported variations do not allow determining actual effect sizes. Thus, the magnitude of relevant changes has to be considered to further support the selection of a small effect size. Because dementia is characterized by rapid progression linked to multiple impairments, it is assumed that even small effects are relevant. The calculation of sample size results in a required sample size of 100 participants for each group (total sample size of 200 participants). Considering reasons for dropout, the sample is set to 405 participants.

5.3.3.6 Dropout

We assume 3 reasons for dropout: (1) withdrawal from the study, (2) missing data, and (3) low attendance or adherence to MEP. Possible reasons for withdrawal are death, hospitalization, serious deterioration in state of health, refusal to participate, etc. Based on the literature review of Blankevoort et al. (2010), a dropout rate of 20 % caused by withdrawal is expected. Missing data occur if participants are not able to complete the entire test battery because of motivational aspects or multiple motor and cognitive impairments. In addition, some participants will not participate at all in postassessments because of illness or other appointments. We assume a missing data rate of 15 %. A total target number of 200 participants (100 per group) for the analysis and an assumed dropout rate (withdrawal and missing data) of 35 % requires enrolling 270 participants into the study. Unfortunately, attendance and adherence are often not stated in previous studies (Forbes et al., 2015). Hence, we decided to double the sample of IG to ensure the required sample of 100 participants in this group. Low attendance and adherence may be caused by illness, motivation, other appointments, disinterest, or other reasons. Hence, a total sample size of 405 participants is required.

All participants will be asked at least twice if they are willing to participate in the assessments to reduce missing data. A familiar caregiver is asked to invite the participant if appropriate. If participants are not willing to complete all measures they are offered to choose assessments they are willing to complete. Moreover, all possible participants will be included in the data collection regardless of whether they discontinued or deviate from the intervention protocol. Caregivers will be asked to support the participants

to get to training sessions to improve attendance. If participants miss a session, they are personally invited to the next training session.

5.3.3.7 Allocation

Group allocation to IG and CG will be performed by minimization to obtain randomized groups with minimum group differences. Subjects rather than care facilities will be randomized to avoid confounding effects of the geographic location, and minimization will be done separately for each care facility based on the baseline criteria MMSE, sex, age, and baseline performance of modified 30-second chair-stand test. Minimization will be performed with the program MinimPy version 0.3 (Saghaei & Saghaei, 2011), which includes a random element. The first participant is allocated randomly to IG or CG. Subsequent participants are allocated to each group correspondingly to achieve the least imbalance between groups. Including a random element, participants will be allocated to the better fitting group with a probability of 70 %. An allocation ratio of 2:1 is selected because of above-mentioned assumptions regarding dropouts. The input order of participants for allocation will be randomly defined by an assigned number for each participant given prior to minimization.

5.3.3.8 Blinding and Pseudonymization

Investigators will be blinded to allocation wherever possible. It is not possible to blind participants or employees of care facilities regarding group allocation.

All data is stored in a strictly pseudonymous form. This is achieved by separating personally identifiable information of participants from data collected during baseline and postassessments. Collation of data is only possible with considerable effort at any time of the study. Thus, individual confidentiality will be ensured before, during, and after the study. Only selected team members have access to coded data.

5.3.3.9 Statistical Analysis

All statistical analysis will be done with SPSS version 23 (IBM Corp). Trained and experienced investigators will evaluate and enter data. Investigators evaluating and entering data are not the same as investigators assessing outcomes. The number of investigators is limited to 2 per assessment method. Prior to actual analysis, interrater reliability (Cohen kappa, Jacob Cohen, 1960; intraclass correlation coefficient, Shrout

& Fleiss, 1979) will be calculated and plausibility (eg, considering range and distribution) will be checked to minimize errors caused by data evaluation and entry.

Because of expected large dropout rate, which can lead to a critical amount of missing data, 2 separate analysis sets are planned: an intention-to-treat analysis and a per-protocol analysis. In the intention-to-treat analysis, all randomized participants regardless of protocol adherence will be included and missing data will be substituted by multiple imputation. Participants with sufficient attendance and adherence to the intervention as well as complete assessments of primary outcomes will be included in the per-protocol analysis, where missing data will not be considered.

Baseline values of participant characteristics will be compared between IG and CG using chi-square tests for categorical data, Mann-Whitney-U tests for nonparametric variables, and t tests for continuous and normally distributed parameters. For all normally distributed data (checked by Shapiro-Wilk test), mean and standard deviation will be calculated, and medians and interpercentile ranges will be calculated for not normally distributed data. Treatment effects will be analyzed using 2-factor ANOVA with repeated measurement. A 2-sided P value less or equal to .05 will be considered to indicate statistical significance. In addition, 95 % confidence intervals and partial Eta^2 will be calculated. Changes in motor and cognitive function are possible mediators and moderators. These mediating and moderating effects on primary outcomes will be analyzed using multiple linear regression models. Additional explorative data analysis exceeding the proposed planned analyses will be performed. Depending on data structure, adequate analysis methods will be defined. These analyses aim to consider further influencing factors or subgroup analysis as well as the development of forecast models.

5.3.4 Results

Enrollment for the study started in May 2015. It is planned to complete postassessments by the beginning of 2017. Results are expected to be available in the first half of 2017.

5.3.5 Discussion

5.3.5.1 Summary

Previous studies have discussed the use of physical activity as additional therapy strategy, and predominately positive effects have been reported. However, the results of these studies are not consistent and they have several methodological limitations. With respect to these limitations, the current study has been carefully designed and thus reflects the following strengths.

5.3.5.2 Strengths

The overall strength is the strong effort to conduct a high-quality trial characterized by a standardized study design, theoretical considerations, an intervention specially designed for IWD, assessments adequate for IWD, a large sample size, and detailed and accurate reporting of methods according to the CONSORT (Boutron et al., 2008; Moher et al., 2010; Schulz et al., 2010) and SPIRIT (Chan, Tetzlaff, Altman et al., 2013; Chan, Tetzlaff, Gøtzsche et al., 2013) statements.

The MEP, which is characterized through dementia-specific methodology and a combination of motor and cognitive tasks, is a major strength of this study. Because of its theoretical foundation and based on primary recommendations of the review by Scharpf et al. (2013), initial guidelines for designing physical activity interventions for IWD can be derived if results support efficiency.

Bearing in mind that most motor assessments are not developed for IWD and their psychometric properties have hardly been systematically established in this specific population (Blankevoort et al., 2013; Bossers et al., 2012), we took several efforts to construct an adequate test battery considering all relevant primary and secondary outcomes. The international expert panel with members from different disciplines where we have discussed possible and adequate measurements as well as general information on performing cognitive and motor measurements in IWD has been an important attempt to enhance quality. In comparison to previous studies, the large sample size is an outstanding feature of this study. To the best of our knowledge, there is no other study with a comparable sample size. Based on studies analyzed for the Cochrane review (Forbes et al., 2015), sample sizes vary between 12 and 148 participants.

This study is designed as a multicenter trial with a sustainable intervention close to everyday life. For instance, the MEP is established on everyday activities such as getting up, walking, or picking things up (see Multimedia Appendix 1). Performing a field study reflects reality in participating care facilities, and results can be more easily transferred to daily routine. Considering sustainability is an important concern of this study and we intend to continue physical activity interventions after study is finished. Thus, employees of care facilities will be educated to guide the MEP. Furthermore, this approach ensures the opportunity for CG to participate in the MEP, which is an important ethical aspect.

5.3.5.3 Challenges

There are several challenges in performing intervention studies in IWD. These are related to the selected study design as well as its target group and thus cannot be avoided. However, it is important to deal with these challenges to minimize their impact.

A big challenge in performing intervention studies with IWD is maintaining blinding to group allocation. Although all investigators will be blinded to group allocation, there is a potential risk that participants will disclose their group allocation during assessments. To minimize this risk, investigators will be asked not to talk about the intervention during assessments.

Working with IWD entails several general challenges as they are often suffering from frailty and multimorbidity. According to different motor and cognitive impairments in IWD, it is not possible to develop an intervention completely suitable for all participants. Hence, some adaptions of the intervention cannot be avoided. However, instructors are asked to minimize such adaptions and adhere to the manual as strictly as possible. Besides this, IWD are vulnerable in relation to attendance, adherence, and missing data. For instance, multiple motor and cognitive impairments partially prevent IWD from participating in all subassessments. Thus, attempts to enhance attendance and adherence as personal communication, support, or repeated invitation are planned.

Further challenges are seen in cooperation with care facilities. Employees assume important responsibilities, such as suggesting potential participants, assessing ADL and state of health, or supporting assessments and intervention. Restricted time or missing expertise is a potential risk for limitations. To reduce such limitations, employees will

be provided detailed information on how to report required data and support for further problems.

5.3.5.4 Implications and Perspectives

Findings of this study will be disseminated through publications and presentations (including information about important protocol modifications). Improving the defined primary outcomes is highly relevant considering the consequences of dementia-related motor deficits as stated in the introduction (American Geriatrics Society et al., 2001; Guralnik et al., 1995; Yümin et al., 2011). Insufficient amounts of physical activity also expedite existing motor and functional impairments in IWD (Brach et al., 2003; Hauer et al., 2006). Therefore, developing adequate physical activity interventions for IWD and offering guidelines is essential. We plan on publishing the MEP and communicating the underlying didactic concept of the training in a detailed manual if it proves to be effective.

This study will contribute to enhance scientific evidence and takes a first look at relations between motor and cognitive skills in IWD. The findings can also be directive for further investigations in the field of prevention, diagnosis, and therapy of dementia.

5.3.5.5 Conclusions

There is a clear need for high-quality studies investigating the effectiveness of physical activity on motor and cognitive skills in IWD. Our study is mainly characterized by a dementia-specific intervention based on scientific knowledge, the combination of motor and cognitive tasks, and a large sample. Findings are highly relevant to influence the multiple motor and cognitive impairments of IWD often participating in limited physical activity. If the MEP proves to be effective, positive influences on everyday life are expected justifying its permanent implementation in care facilities.

Abbreviations

ADAS-Cog: Alzheimer Disease Assessment Scale–Cognitive Subscale; ADL: activities of daily living; ANOVA: analysis of variance; CERAD: Consortium to Establish a Registry for Alzheimer's Disease; CG: control group; CONSORT: Consolidated Standards of Reporting Trials; E-ADL-Test: Erlangen Test of Activities of Daily Living; FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques 4; ICD-10: International Statistical Classification of Diseases and Related Health Problems, Tenth Edition; IG: intervention group; IWD: individuals with dementia; MEP: multimodal exercise program; MMSE:

Mini-Mental State Examination; SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

Declarations

Acknowledgments

All authors are members of the research team and participated in the implementation of the study. AW conceived the idea of this study along with ST and AS. All authors contributed to the conception and design of the study. ST coordinates the study under direct supervision of AW. ST and AS are responsible for the implementation of the study. AS developed and supervises the multimodal exercise program. All authors were involved in planning and writing the study protocol. ST and BB wrote the study protocol. AS and CN helped draft this manuscript. All authors provided critical feedback and approved the final manuscript.

This project is financially supported by the Dietmar Hopp Stiftung (St. Leon-Rot, Germany). The sponsor does not have any role in the design of the study; its execution, data collection, analysis, or interpretation; the decision to submit results; or the writing of the report.

Ethical approval has been obtained from the ethics commission of the Karlsruhe Institute of Technology (11.03.2015, reference number: 7712.14-0508-0). An informed consent will be signed by participants or their legal guardians prior to the study.

We are grateful to all participating in the study and to our research stuff. We thank Luisa Appelles for the coordination of fieldwork, Dr Annegret Mündermann for her writing assistance on behalf of the authors, Dr Doris Oriwol for the statistical consulting, and all participants of the expert panel for developing the test battery. We acknowledge support by Deutsche Forschungsgemeinschaft and the Open Access Publishing Fund of Karlsruhe Institute of Technology.

Conflicts of Interest

None declared.

Multimedia Appendix

Multimedia Appendix 1. Sample session of the multimodal exercise program.

Multimedia Appendix 2. Description of the assessments.

References

All references of manuscript III are included in the List of References at the end of this thesis.

6 Performing a randomized controlled trial to investigate the effectiveness of physical activity in individuals with dementia

6.1 Effectiveness of a dementia-specific multimodal exercise program on motor performance in individuals with dementia

Manuscript IV

Summary: Physical activity gains in importance as a therapy strategy to improve motor performance in IWD (Ahlskog et al., 2011). Despite various investigations, evidence of the effectiveness of physical activity on motor performance in IWD cannot be ensured. Recent reviews refer to methodologic limitations, risk of bias, as well as lack of comparability of previous studies, and thus indicate the need for further high-quality studies (Blankevoort et al., 2010; Hernández et al., 2015; Lam, Huang et al., 2018). Aiming to contribute to enhancing evidence, manuscript IV reports findings of a high-quality RCT investigating the effectiveness of physical activity on motor performance in IWD. It refers to research question B1 and hypothesis 1.

Version: This is the author's original before peer review.

Publication status: submitted

Reference:

Barisch-Fritz, B., Trautwein, S., Scharpf, A., Krell-Roesch, J., & Woll, A. (submitted). Effects of a 16-week multimodal exercise program on motor performance in individuals with dementia: A multicenter randomized controlled trial.

Effects of a 16-Week Multimodal Exercise Program on Motor Performance in Individuals with Dementia: A Multicenter Randomized Controlled Trial[12]

Bettina Barisch-Fritz[1], PhD; Sandra Trautwein[1]; Andrea Scharpf[1]; Janina Krell-Roesch[1,2], PhD; and Alexander Woll[1], PhD, Prof.

[1] Karlsruhe Institute of Technology, Institute of Sports and Sports Science, Karlsruhe, Germany

[2] Mayo Clinic, Translational Neuroscience and Aging Lab, Scottsdale, Arizona, USA

6.1.1 Abstract

Background and Purpose: Dementia is not restricted to cognitive impairments but also affects physical performance. In individuals with dementia (IWD), a decline in physical performance directly impairs independent life and quality of life increasing with disease progression. As there is no conclusive evidence of the effectiveness of physical activity interventions on physical performance of IWD, examining factors that potentially preserve physical performance is of high importance. Reasons for missing evidence are small numbers of high-quality studies, large heterogeneity in used methods and insufficient reporting of methods. The aim of this study was to investigate effects of a 16-weeks multimodal exercise program (MEP) combining motor and cognitive tasks on physical performance in IWD. Additionally, we identified characteristics of responders to MEP.

Methods: A multicenter randomized controlled trial with appropriate assessment methods and a standardized MEP adjusted to IWD was conducted. We included 319 IWD of mild to moderate severity, aged over 65 years. At baseline and after MEP, we assessed physical performance with the primary outcomes: mobility, balance, and strength and function of lower extremities. Potential effects were identified by using two-factor analyses of variance with repeated measurements within two samples, i.e. intention-to-treat and per protocol sample. Additionally, we compared characteristics related to physical performance between positive, non-, and negative responders.

Results and Discussion: Both analysis procedures did not reveal statistically significant time*group effects. We identified 56-66 % non- or positive responders. Furthermore, positive responders in balance and strength and function of lower extremities had statistically significant lower baseline performance in the same variable. Overall, effects

[12] Some minor formal adaptions were made to the version of manuscript IV presented in this thesis to ensure uniform formatting.

of MEP on physical performance were not statistically significant which possibly resulted from the high heterogeneity of the sample. In addition, we observed a trend for IWD with lower physical performance at baseline benefitting more than those with higher baseline performance.

Conclusions: We recommend a higher degree of individualization which might improve overall effectiveness. These findings support research focusing on physical activity and dementia, particularly on designing and implementing exercise interventions.

Trial registration: DRKS00010538 (German Clinical Trial Register, date of registration: 01 June 2016, retrospectively registered).

Keywords: Physical activity, dementia, physical performance, cognitive impacts, motor impacts

6.1.2 Introduction

The successful completion of physically demanding tasks mainly requires the basic abilities: mobility, balance, and strength. These basic abilities are summarized by the construct of “physical performance” (Rydwik, Frändin, & Akner, 2004). A decline in physical performance, as caused by normal aging processes or various disorders, reduces the capacity to successfully perform activities of daily living (e.g. eating, bathing) which in turn leads to an increased dependency in everyday life (Martyr & Clare, 2012; Wennie Huang, Perera, VanSwearingen, & Studenski, 2010). Maintenance of independency and autonomy of individuals makes a significant contribution to an individual’s quality of life (Telenius et al., 2013; Vermeulen et al., 2011; Yümin et al., 2011).

A decline in mobility and balance and consequently activities of daily living are particularly common in individuals with dementia (IWD) (Allan et al., 2005; Manckoundia et al., 2006; van Iersel et al., 2004). The symptoms of dementia – a syndrome of several different types of usually chronic and progressive diseases of the brain (World Health Organization, 2016) – are not restricted to cognitive skills but also influence physical performance (Kido et al., 2010; Leandri et al., 2009). Compared to cognitively unimpaired older adults and also individuals with mild cognitive impairments, IWD show a poorer performance in balance (Kato-Narita et al., 2011) and mobility (Eggermont et al., 2010; Gras et al., 2015). Concerning gait, especially walking speed is decreased, stride length shortened, and double support time increased (Coelho et al., 2013; Ijmker

& Lamoth, 2012). Due to infrequent investigation, information about strength and endurance is somewhat conflicting but overall also indicative of poorer performance (Arshinta et al., 2018; Manckoundia et al., 2006). Consequently, supervision of expenditures and burdens for the health care system are high (Alzheimer's Disease International, 2015; Livingston et al., 2017; Schulze et al., 2015). Especially given that the number of IWD is currently at 50 million worldwide (Alzheimer's Disease International, 2018) and expected to further increase to more than 152 million by the year 2050 (Alzheimer's Disease International, 2018), the World Health Organization has declared dementia a public health priority (World Health Organization, 2012).

Even though cognitive impairments are considered the hallmark of dementia (Waldemar et al., 2007), deficits in physical performance as further symptoms of dementia are getting more and more attention even in clinical diagnostics. In particular, difficulties in walking which requires complex cognitive and motor processes can predict an increased risk of developing cognitive deficits in older adults (Bridenbaugh & Kressig, 2015; Buracchio, Dodge, Howieson, Wasserman, & Kaye, 2010; Verghese et al., 2007). The best predictors for dementia are both mild cognitive impairments as well as slower gait, and this combination was first described as concept of cognitive risk syndrome by Verghese et al. (2014).

It has been postulated that there are causal relations between motor decline and cognitive impairments (Buchman & Bennett, 2011), particularly as the pathological changes underlying Alzheimer's disease also affect regions known to subserve physical performance (Burns, Galvin, Roe, Morris, & McKeel, 2005; Schneider et al., 2006). It might be more obvious given that control systems for motor regulation, initiation, planning, and execution are located in several cortical and subcortical regions (Fogassi & Luppino, 2005; Halsband & Lange, 2006; Lehéricy et al., 2006; Rizzolatti & Luppino, 2001). Furthermore, it is generally hypothesized that motor impairments increase with disease progression (Bridenbaugh & Kressig, 2015) and are also accompanied by a higher risk of falls (Allali et al., 2016; Amboni et al., 2013). In light of these close associations between motor function and cognitive performance, it is paramount to investigate whether a physical activity intervention has an impact on physical performance and thus motor impairments in IWD.

In the absence of a cure for dementia and considerable side effects of currently prescribed drugs (Hugo & Ganguli, 2014; Versijpt, 2014), non-pharmacological interventions have gained increasing importance in recent years. Particularly, physical activity interventions may be beneficial for IWD. Indeed, there is growing evidence that physical activity interventions may slow the decline in physical performance (Blankevoort et al., 2010). However, in IWD this evidence is not as clear as in institutionalized elderly persons with multiple diagnoses, where strong evidence of a positive effect of physical activity on muscle strength and mobility was found (Rydwik et al., 2004). Even if reviews mainly report a positive effect of physical activity on lower extremity strength and activities of daily living in IWD (Blankevoort et al., 2010; Forbes et al., 2015; Heyn et al., 2008; Littbrand et al., 2011; Pitkälä, Savikko et al., 2013; Potter et al., 2011; Suttanon et al., 2010), effect sizes for balance and mobility however vary from small negative to large values but are in general higher than for other motor skills (Blankevoort et al., 2010; Suttanon et al., 2010). Therefore, a conclusive evidence of the impact of physical activity interventions on physical performance does not exist. This may be due to small numbers of high-quality studies and large heterogeneity in used methods as well as insufficient reporting of methods particularly with regard to intervention modalities (e.g. detailed training parameters; Blankevoort et al., 2010; Forbes et al., 2015; Scharpf et al., 2013).

The aim of this study is to investigate the effects of a multimodal exercise program (MEP) that combines physical and cognitive tasks on the primary outcomes of physical performance in IWD. The secondary aim is to overcome existing methodological deficits by the design and conduct of a multicenter randomized controlled trial (RCT) with appropriate assessment methods as well as a sustainable intervention adjusted to the characteristics of this sample. Therefore, the central hypothesis is that physical performance of IWD changes after a 16-weeks MEP and as compared to a control condition consisting of conventional treatment. Additionally, we want to display the individual responses to MEP as recommended by the FDA (U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, & U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health, 2006). The responder analysis allows for assessing clinical relevance and helps to increase the translation of the results into clinical practice.

6.1.3 Methods

The details of this study are comprehensively described in the study protocol (Trautwein et al., 2017) following the guidelines and recommendations of the Consolidated Standards of Reporting Trials (CONSORT) statements (Moher et al., 2010; Schulz et al., 2010). This section provides a shorter description of the methods referring to the CONSORT statements. The study was retrospectively registered in the German National Register of Clinical Trials (DRKS00010538) and was approved by the ethics committee of the Karlsruhe Institute of Technology (Karlsruhe, Germany).

6.1.3.1 Study design and participants

The design of the study to address the central hypothesis includes a multicenter parallel-group RCT with baseline and post assessments and an allocation ratio of 2:1 for an intervention (IG) and control group (CG), respectively.

A power analysis (G*Power 3, Version 3.1.9.2, Faul et al., 2007, two-factor analysis of variance [ANOVA] with repeated measurement, two groups and two measurements, $\alpha=0.05$, $1-\beta=0.80$, $\eta^2=0.01$) determined a required total sample size of 200 participants. The calculated sample size is based on the assumptions that even small effects are relevant with regard to the rapid disease progression seen in IWD.

Recruitment for this study took place in 36 care facilities in South-Western Germany. Employees of care facilities were asked to identify possible participants. Before entering the study, written consent of participants or their legal guardians was reclaimed. After baseline assessment, final inclusion or exclusion was determined according the following criteria:

Inclusion criteria: a) diagnosis of dementia or “suspected” dementia; b) Alzheimer disease, vascular dementia or other primary dementia; c) mild to moderate stage of dementia (Mini-Mental State Examination [MMSE]: 10-24); d) age above 65 years; e) walking ability of about ten meters with or without walking aid; and f) clearance from general practitioner.

Exclusion criteria: a) secondary dementia; b) other severe cognitive impairments; c) other severe neurological disease; d) other severely acute diseases; and e) severe motor impairments.

6.1.3.2 Randomization and intervention

After baseline assessment, participants were randomly allocated to IG or CG by minimization (MinimPy Version 0.3). Randomization was done stratified by each care facility with an allocation ratio of 2:1 in favor of the exercise arm. Investigators were blinded where possible.

The intervention program was implemented in conventional treatment (individualized medication, care, therapeutic applications etc.) of participants of the IG. Participants of the CG solely continued their conventional treatment. The MEP is a specifically developed intervention that combines physical and cognitive tasks considering the special needs of IWD. It was delivered two times a week on non-consecutive days for 16 weeks by two skilled instructors in a group setting of up to 12 participants. Each session lasted 60 minutes, with training time being about 45 minutes.

Beside other didactic aspects like adaptation to cognitive levels of participants or adjusted communication, the MEP aimed at giving the participants orientation and familiarity as well as a sense of security through ritualization. To this end, each session was similarly structured i.e. the three parts arrival, destination, and departure were embedded in experienced journeys.

Main physical contents over the entire intervention period can be divided into tasks for strength (43 %), balance (25 %), endurance (16 %), flexibility (13 %), and not further specified tasks (3 %). These tasks were carried out with medium to submaximal intensity. Cognitive tasks focused on the stimulation of memory, attention, language, and executive functions. Intensity of physical and cognitive tasks was enhanced throughout the intervention period.

6.1.3.3 Outcomes and assessments

Primary outcomes were motor qualities mobility, balance, as well as strength and function of lower extremities. These motor skills were assessed by means of several assessments listed in Table 22.

Table 22. Primary outcome parameters

Primary outcome	**Assessments (at baseline and after 16-weeks multimodal exercise program)**
Mobility	6-meter walk test, 6m WT(Graham et al., 2008)
	Timed Up & Go Test, TUG (Podsiadlo & Richardson, 1991)
Balance	Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, FICSIT-4 (Rossiter-Fornoff et al., 1995)
Strength and function of lower extremities	Modified 30-second chair stand test, modified 30s CST (Blankevoort et al., 2013; Jones et al., 1999)
	Modified Short Physical Performance Battery, modified SPPB (Guralnik, Simonsick et al., 1994)

The used motor assessments were finalized after an international expert panel focusing on appropriateness as well as feasibility. Details on this discussion and the final assessment battery can be found in Trautwein, Barisch-Fritz et al. (2019). The baseline and post assessments were executed within the care facilities by trained study staff.

Mobility was assessed using 6-meter walk test (6m WT) (Graham et al., 2008) and Timed Up & Go Test (TUG) (Podsiadlo & Richardson, 1991). 6m WT was captured for two times. Participants were not explicitly asked to walk the marked distance of six meters, to reduce bias by testing situation. TUG was conducted twice by asking participants to rise from a chair, walk three meters, turn around, go back and sit down again. Walking aids were allowed for all assessments of mobility.

Static balance was determined by Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4 (FICSIT-4) where participants were asked to perform different standing positions for ten seconds. The four tasks with rising complexity are Romberg, semi tandem, tandem, and single leg. The performance is rated with zero to five points according to number and time of finished positions (Rossiter-Fornoff et al., 1995).

Strength and function of lower extremities was assessed by modified 30-second chair stand test (modified 30s CST) and the modified Short Physical Performance Battery (modified SPPB). For modified 30s CST, participants were asked to stand up from a chair (height 46 centimeters, with armrests) as often as possible during 30 seconds. Time to perform five repetitions was recorded. The modified version allows to use the armrests (Blankevoort et al., 2013; Jones et al., 1999) which is essential for the majority of elderly IWD. The Short Physical Performance Battery (SPPB) (Guralnik, Simonsick et al., 1994) focusses on the function of the lower extremities. Being based on three

subtests of mobility, standing balance, and rising from a chair, it is a reliable and valid measure in community dwelling older adults (Freiberger et al., 2012; Mijnarends et al., 2013). In IWD living in care facilities, a pilot study concluded that relative reliability is acceptable whereas absolute reliability is rather moderate (Fox et al., 2014). Modified SPPB consists like SPPB of the three subscales walking speed for mobility, FICSIT-4 for standing balance, and time for five repetitions of modified 30s CST with arm use for strength as well as function of the lower extremities. The scoring comprises a scale of 0-12, with a better function as higher the score is (Olsen & Bergland, 2017).

The secondary outcomes were overall cognition determined by MMSE (Folstein et al., 1975). Moreover, body mass and height were measured. Body Mass Index (BMI) was calculated on the base of this information. Adherence of participants was documented by instructors for each session. Sample characteristics or potential cofounders such as age, sex, use of walking aids, diagnosis and etiology of dementia, as well as number of medications were assessed by questionnaire before baseline assessments. This questionnaire as well as the Cumulative Illness Rating Scale (CIRS) (Linn et al., 1968) was completed for each participant by employees of each care facility or the general practitioner. Participants or their legal guardian signed a written consent prior to participation in the study.

6.1.3.4 Statistical analysis

The statistical analysis to investigate changes due to the MEP between IC and CG were calculated within two samples. In an intention-to-treat analysis, all participants that fulfilled the inclusion criteria and were randomized to IG or CG were considered, except for deceased participants. For this reason, multiple imputation procedure (fully conditional specification imputation method, ten imputations, and ten iterations) of the primary outcomes (6m WT, TUG, FICSIT-4, modified 30s CST, modified SPPB) was used to handle missing data. For the multiple imputation several variables were considered as predictors such as all primary outcomes supplemented by adherence as well as related motor and cognitive performance. To ensure plausibility of imputed data further constraints were defined like minimum and maximum values according to observed range in each variable, rounding according to original data, 100 maximal case draws, and ten maximal parameter draws. As final estimates, we considered pooled

results as provided by SPSS or reported ranges observed throughout the imputations, if SPSS did not support the pooling procedure.

Within a second sample resulting from the per protocol analysis, we pursued the same goals. However, the sample was smaller as we assumed an adherence to the MEP of at least 75 % reflecting an active participation in at least 24 of 32 training sessions. Further differences in sample size resulted from individual participation in the several assessments, where missing data were not considered. This participation is highly dependent on mood and form of the day of each IWD and led to several challenges for the investigators.

Baseline values of participant characteristics were compared between IG and CG using Chi-square tests for categorical data, Mann-Whitney-U-Tests for non-parametric variables, and T-tests for continuous and normally distributed parameters. Normally distributed data (checked by Shapiro-Wilks-Test and relevant graphs) are presented by means and standard deviations. Treatment effects were analyzed and presented as within group effects (differences from baseline to post assessments) and time*group effects (changes from baseline to post assessments between groups). Therefore, paired T-Tests and two-factor ANOVA with repeated measurement were calculated for the primary outcomes.

The responder analysis is based on distribution-based methods i.e. information about the standard error of measurement of the assessments within the per protocol analysis. To this end, a positive change from baseline of 10 % and more is defined as positive responder. Changes in between positive and negative 10 % are defined as non-responder and negative changes of 10 % or more are define as negative responders. Selected sample characteristics were compared between positive, non-, and negative responders using Kruskal-Wallis-Tests and one-factor ANOVA. For post-hoc analyses, we used Dunn-Bonferroni-Tests and Tukey-Kramer post-hoc tests, respectively. R and partial Eta² served as effect sizes. All statistical analyses were done with IBM SPSS Version 25 (IBM Corporation, Armonk, USA). The significance level was set for all tests at $p<0.05$. Evaluation and entering of data were conducted by trained and experienced investigators.

6.1.4 Results

6.1.4.1 Sample characteristics

Between March 2015 and March 2017, 600 IWD were screened for eligibility. Out of this sample, we considered 319 persons suitable for the study. After examination of the baseline assessments, 201 participants were allocated to IG and 118 to CG. The overall dropout rate was 8 %. Figure 8 shows the overall flow of participants.

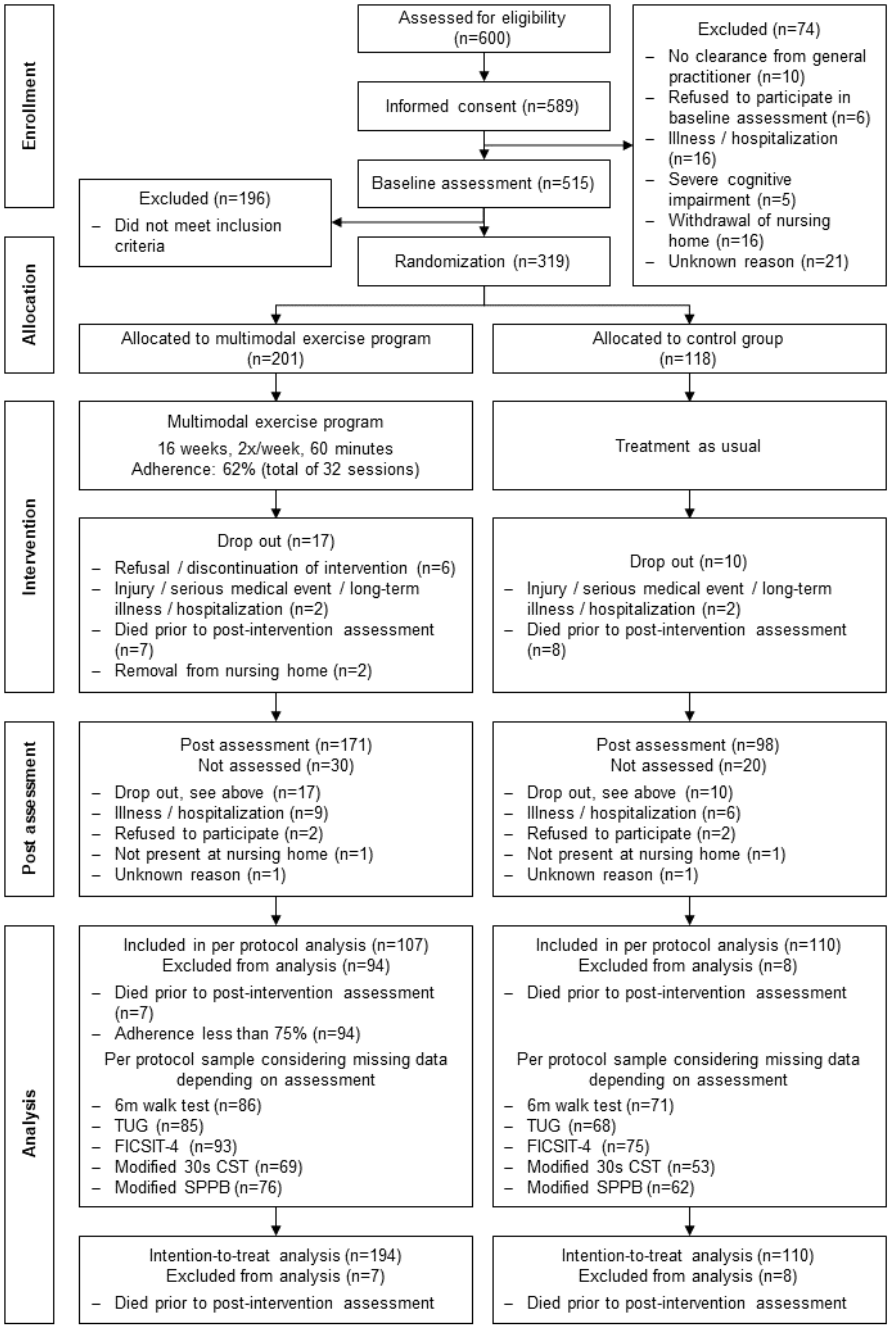


Figure 8. Flow of participants (6m WT: 6-meter walk test, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, modified 30s CST: modified 30-second chair stand test, modified SPPB: Short Physical Performance Battery, n: number of individuals with dementia, TUG: Timed Up & Go Test).

Due to dropouts caused by death a final sample size of 304 participants was available for the intention-to-treat analysis (Table 23). The sample for the per protocol analysis was further reduced by 87 participants within the IG due to less than 75 % adherence on the MEP (Table 24).

Table 23. Sample characteristics of participants at baseline by intention-to-treat analysis

Characteristics	**Total sample [n=304]**	**Intervention group [n=194]**	**Control group [n=110]**	**Group differences**
	M (SD)	**M (SD)**	**M (SD)**	**T(df)/ Chi²(df), p**
Age (years)	86.1 (6.1)	85.8 (6.3)	86.6 (5.8)	t(302)=1.135, p=0.257
MMSE	17.0 (4.1)	16.9 (4.3)	17.1 (3.8)	t(250.853)=0.389, p=0.698
BMI (kg/m²) (n=270)	28.0 (4.7)	28.5 (4.7)	27.2 (4.8)	**t(268)=-2.307, p=0.022**
CIRS (n=178)				
Morbidity Index	9.3 (4.8)	9.2 (4.4)	9.5 (5.6)	t(176)=0.469, p=0.640
Severity Index	1.6 (0.4)	1.6 (0.4)	1.6 (0.4)	t(176)=0.024, p=0.981
Number of medications (n=234)	6.9 (3.9)	7.5 (3.8)	6.0 (4.0)	**t(232)=-2.686, p=0.008**
	n (%)	**n (%)**	**n (%)**	
Sex				Chi²(1)=1.223, p=0.269
Female	262 (86.2)	164 (84.5)	98 (89.1)	
Male	42 (13.8)	30 (15.5)	12 (10.9)	
Diagnosis of dementia				Chi²(2)=3.693, p=0.158
yes	200 (65.8)	129 (66.5)	71 (64.5)	
no	55 (18.1)	39 (20.1)	16 (14.5)	
unknown	49 (16.1)	26 (13.4)	23 (20.9)	
Type of dementia				Chi²=9.005, p=0.050
Alzheimer's disease	51 (16.8)	36 (18.6)	15 (13.6)	
Vascular dementia	45 (14.8)	34 (17.5)	11 (10.0)	
Mixed dementia	8 (2.6)	4 (2.1)	4 (3.6)	
other	4 (1.3)	4 (2.1)	0 (0.0)	
unknown	92 (30.3)	51 (26.3)	41 (37.3)	
no/unknown diagnosis	104 (34.2)	65 (33.5)	39 (35.5)	
Use of walking aid				Chi²(2)=4.104, p=0.128
no walking aid	64 (21.5)	46 (24.5)	18 (16.5)	
walker	216 (72.7)	134 (71.3)	82 (75.2)	
waking stick/s	17 (5.7)	8 (4.3)	9 (8.3)	

BMI: Body Mass Index, CIRS: Cumulative Illness Rating Scale, df: degree of freedom, M: mean, MMSE: Mini-Mental State Examination, n: number of individuals with dementia, SD: standard deviation

Table 24. Sample characteristics of participants at baseline by per protocol analysis

Characteristics	**Total sample [n=217]**	**Intervention group [n=107]**	**Control group [n=110]**	**Group differences**
	M (SD)	**M (SD)**	**M (SD)**	**T(df)/ Chi²(df), p**
Age (years)	85.9 (6.3)	85.2 (6.7)	86.6 (5.8)	t(215)=1.617, p=0.107
MMSE	16.9 (4.1)	16.8 (4.4)	17.1 (3.8)	t(208.706)=0.603, p=0.547
BMI (kg/m²) (n=199)	27.8 (4.7)	28.4 (4.5)	27.2 (4.8)	t(197)=-1.960, p=0.051
CIRS (n=124)				
Morbidity Index	9.3 (5.0)	9.0 (4.4)	9.5 (5.6)	t(122)=0.554, p=0.581
Severity Index	1.6 (0.4)	1.5 (0.4)	1.6 (0.4)	t(122)=0.835, p=0.405
Number of medications (n=167)	6.8 (4.0)	7.5 (3.9)	6.0 (4.0)	**t(165)=-2.347, p=0.020**
	n (%)	**n (%)**	**n (%)**	
Sex				Chi²(1)=1.161, p=0.281
Female	188 (86.6)	90 (84.1)	98 (89.1)	
Male	29 (13.4)	17 (15.9)	12 (10.9)	
Diagnosis of dementia				Chi²(2)=4.154, p=0.125
yes	145 (66.8)	74 (69.2)	71 (64.5)	
no	37 (17.1)	21 (19.6)	16 (14.5)	
unknown	35 (16.1)	12 (11.2)	23 (20.9)	
Type of dementia				Chi²(4)=6.563, p=0.134
Alzheimer's disease	29 (13.4)	14 (13.1)	15 (13.6)	
Vascular dementia	33 (15.2)	22 (20.6)	11 (10.0)	
Mixed dementia	6 (2.8)	2. (1.9)	4 (3.6)	
other	2 (0.9)	2. (1.9)	0 (0.0)	
unknown	75 (34.6)	34 (31.8)	41 (37.3)	
no/unknown diagnosis	72 (33.2)	33 (30.8)	39 (35.5)	
Use of walking aid				Chi²(2)=4.674, p=0.097
no walking aid	48 (22.5)	30 (28.8)	18 (16.5)	
walker	148 (69.5)	66 (63.5)	82 (75.2)	
waking stick/s	17 (8.0)	8 (7.7)	9 (8.3)	

BMI: Body Mass Index, CIRS: Cumulative Illness Rating Scale, df: degree of freedom, M: mean, MMSE: Mini-Mental State Examination, n: number of individuals with dementia, SD: standard deviation

The sample characteristics of the two samples based on the two analysis procedures can be found in Table 23 and Table 24. The mean (SD) age of the participants within the intention-to-treat analysis was 86.1 (6.1) with 86.2 % female out of 304 participants. Within the per protocol analysis, the mean (SD) age was 85.9 (6.3) with 86.6 % female out of 217 participants. MMSE ranged within both samples from 10 to 24 points with mean (SD) values in the intention-to-treat analysis of 17.0 (4.1) and the per protocol

analysis of 16.9 (4.1) which is indicative of a mild to moderate severity of dementia at baseline assessment. The CIRS morbidity index was on average 9.3 (4.8) in the intention-to-treat analysis and 9.3 (5.0) in the per protocol analysis. The number of medications included a mean of 6.9 (3.9) in the intention-to-treat analysis and 6.8 (4.0) in the per protocol analysis. Only 21.5 % of the intention-to-treat and 22.5 % of the per protocol sample needed no walking aid, the others used walkers or walking sticks. These values indicate the presence of comorbidities as well as high frailty of participants. Within both samples, no statistically significant differences at baseline for the variables that characterize IWD between IG and CG were identified, except for the number of medications and BMI.

6.1.4.2 Effects of the multimodal exercise program on motor performance

For the intention-to-treat analysis missing data were identified ranging between 9.1 % and 43.9 %. Beside participants that died during intervention period, several reasons like medical constrains, refusal, discontinuation of the assessment led to not joining individual assessments or the whole post-intervention assessment account for the missing data. Table 25 presents the results on the effects of the MEP on physical performance as derived from the intention-to-treat analysis. Results are presented by baseline and post-intervention values, differences between baseline and post-intervention assessments, group differences at baseline, within group time effects, and time*group effects including effect sizes. Participants of the IG had a mean adherence of 62 %.

The intention-to-treat analysis revealed no statistically significant differences in the primary outcomes between the two groups at baseline as well as within the groups after the 16-weeks MEP. Similarly, we observed no statistically significant time*group effects.

Table 25. Effect of the multimodal exercise program on physical performance of individuals by intention-to-treat analysis

		Baseline IG: n=194 CG: n=110 [M (SE)]	**Group differences at baseline [t(df), p]**	**Post IG: n=194 CG: n=110 [M (SE)]**	**Difference baseline - post [M (SE), [CI_{95}]]**	**Within group time effects [t(df), p]**	**Time*group effects $F(df_{numerator}, df_{denominator})$, p**	**Effect size η_p^2**
6m WT (sec)	IG	11.53 (0.44)	t(8435)=-0.154, p=0.878	10.81 (0.30)	0.72 (0.40), [-0.07, 1.51]	t(199)=1.783, p=0.076*	F(1,302)=0.202 to 3.468, p=0.064 to 0.653 [b]	0.001 to 0.011
	CG	11.43 (0.47)		11.54 (0.44)	-0.10 (0.53), [-1.15, 0.95]	t(39)=-0.196, p=0.846		
TUG (sec)	IG	26.60 (1.23)	t(532)=-1.000, p=0.318	25.75 (1.05)	0.84 (1.26), [-1.66, 3.35]	t(30)=0.673, p=0.506*	F(1,302)=1.183 to 6.232, p=0.013 to 0.278 *, [a, b]	0.004 to 0.020
	CG	24.74 (1.15)		26.54 (1.29)	-1.80 (1.13), [-4.04, 0.44]	t(64)=-1.590, p=0.117*		
FICSIT-4	IG	1.97 (0.10)	t(208)=-0.505, p=0.614	2.06 (0.11)	-0.09 (0.11), [-0.31, 0.13]	t(47)=-0.832, p=0.410	F(1,302)=0.024 to 5.453, p=0.020 to 0.876 *, [a]	0.000 to 0.018
	CG	1.89 (0.12)		1.77 (0.15)	0.11 (0.16), [-0.20, 0.43]	t(32)=0.711, p=0.482		
Modified 30s CST	IG	7.67 (0.29)	t(127)=-0.118, p=0.906	7.41 (0.30)	0.26 (0.30), [-0.33, 0.85]	t(26)=0.878, p=0.388*	F(1,302)=0.006 to 2.281, p=0.132 to 0.939	0.000 to 0.007
	CG	7.62 (0.35)		7.59 (0.36)	0.03 (0.37), [-0.71, 0.76]	t(45)=0.074, p=0.942		
Modified SPPB	IG	5.88 (0.20)	t(1552)=-0.313, p=0.754	6.04 (0.23)	-0.15 (0.18), [-0.52, 0.21]	t(59)=-0.841, p=0.403	F(1,302)=0.758 to 9.541, p=0.002 to 0.385 *, [a, b]	0.003 to 0.031
	CG	5.78 (0.24)		5.43 (0.27)	0.35 (0.25), [-0.14, 0.85]	t(48)=1.412, p=0.165*		

6m WT: 6-meter walk test, CG: control group, CI_{95}: 95 % confidence interval, df: degrees of freedom, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, IG: intervention group, M: mean, modified 30s CST: modified 30-second chair stand test, modified SPPB: modified Short Physical Performance Battery, n: number of individuals with dementia, SE: standard error, TUG: Timed Up & Go Test

* statistically significant in single imputations, [a] variance homogeneity not fulfilled in all imputations, [b] covariance homogeneity not fulfilled in all imputations

Table 26. Effect of the multimodal exercise program on physical performance of individuals with by per protocol analysis

		Baseline [M (SD)]	**Group differences at baseline [t(df), p]**	**Post [M (SD)]**	**Difference baseline - post [M (SD), [CI_{95}]]**	**Within group time effects [t(df), p]**	**Time*group effects $F(df_{numerator}, df_{denominator})$, p**	**Effect size η_p^2**
6m WT (sec)	IG; n=86	9.79 (3.15)	t(155)=-1.620, p=0.107	9.81 (3.52)	-0.02 (2.66), [-0.60, 0.55]	t(85)=-0.085, p=0.933	F(1,155)=0.000, p=0.986 [b]	0.000
	CG; n=71	10.75 (4.28)		10.79 (3.74)	-0.03 (4.17), [-1.02, 0.95]	t(70)=-0.068, p=0.946		
TUG (sec)	IG; n=85	20.84 (9.99)	t(151)=0.919, p=0.360	21.84 (10.64)	-1.01 (7.71), [-2.67, 0.66]	t(84)=-1.201, p=0.233	F(1,151)=0.419, p=0.518	0.003
	CG; n=68	22.32 (9.84)		24.11 (10.71)	-1.79 (7.16), [-3.52, -0.06]	**t(67)=-2.065, p=0.043**		
FICSIT-4	IG; n=93	2.25 (1.44)	t(194.172)=-1.477, p=0.141	2.32 (1.42)	-0.07 (1.42), [-0.36, 0.22]	t(92)=-0.475, p=0.636	F(1,166)=0.328, p=0.567	0.002
	CG; n=75	1.87 (1.20)		1.81 (1.27)	0.05 (1.34), [-0.26, 0.36]	t(74)=0.344, p=0.732		
Modified 30s CST	IG; n=69	8.29 (3.52)	t(120)=-0.374, p=0.709	8.17 (3.46)	0.12 (2.45), [-0.47, 0.71]	t(68)=0.418, p=0.678	F(1,120)=0.302, p=0.584	0.003
	CG; n=53	8.05 (3.59)		8.21 (3.37)	-0.16 (3.25), [-1.06, 0.74]	t(52)=-0.359, p=0.721		
Modified SPPB	IG; n=76	6.78 (2.82)	t(136)=-1.245, p=0.215	6.87 (2.99)	-0.09 (2.16), [-0.59, 0.40]	t(75)=-0.372, p=0.711	F(1,136)=0.980, p=0.324	0.007
	CG; n=62	6.19 (2.63)		5.90 (2.61)	0.29 (2.37), [-0.31, 0.89]	t(61)=0.964, p=0.339		

6m WT: 6-meter walk test, CG: control group, CI_{95}: 95 % confidence interval, df: degrees of freedom, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, IG: intervention group, M: mean, modified 30s CST: modified 30-second chair stand test, modified SPPB: modified Short Physical Performance Battery, n: number of individuals with dementia, SD: standard deviation, TUG: Timed Up & Go Test

[a] variance homogeneity not fulfilled, [b] covariance homogeneity not fulfilled

Statistically significant results appear bold for α=0.05, considering adjusted significance levels using Bonferroni-Holm correction no statistically significant results were observed

The per protocol analysis of the effects of MEP is presented in Table 26 and shows baseline and post-intervention values, differences between baseline and post-intervention assessments, group differences at baseline, within group time effects, and time*group effects including effect sizes. Participants of the IG had a mean adherence of 91 %. Overall, the per protocol analysis revealed no statistically significant differences in the primary outcomes between the two groups at baseline as well as within the groups after the 16-weeks MEP. As in the intention-to-treat analysis no statistically significant time*group effects were observed.

6.1.4.3 Differences in characteristics between positive, negative, and non-responders (intervention group, per protocol analysis)

Between 28 % and 40 % of participants in the IG improved their motor performance by at least 10 % (considered as positive responders). Moreover, physical performance did not change in 26 % to 37 % of participants (considered as non-responders), while 34 % to 44 % showed a decline in motor performance by at least 10 % (considered as negative responders). Table 27 displays the proportion of positive, non-, and negative responders in the IG depending on motor assessment, as well as mean changes in motor performance.

Table 27. Positive, non-, and negative responders in the intervention group and mean changes in physical performance (per protocol analysis)

	All		Negative responders		Non-responders		Positive responders	
	n	Mean change, M (SD) [%]	n [%]	Mean change, M (SD) [%]	n [%]	Mean change, M (SD) [%]	n [%]	Mean change, M (SD) [%]
6m WT	86	0.0 (2.7)	34%	2.9 (1.7)	31%	-0.3 (0.5)	35%	-2.5 (1.7)
TUG	85	1.0 (7.7)	35%	7.9 (6.5)	37%	0 (1.1)	28%	-6.3 (6.5)
FICSIT-4	93	0.1 (1.4)	38%	-1.3 (0.6)	28%	0 (0)	34%	1.6 (0.9)
Modified 30s CST	69	-0.1 (2.5)	44%	-2.3 (1.4)	26%	0.3 (0.5)	30%	2.6 (1.4)
Modified SPPB	76	0.1 (2.2)	34%	-2.2 (1.4)	26%	0.0 (0.6)	40%	2.1 (1.2)

6m WT: 6-meter walk test FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, n: number of individuals with dementia, M: mean, modified 30s CST: 30-second chair stand test, modified SPPB: Short Physical Performance Battery, SD: standard deviation, TUG: Timed Up & Go Test

Positive, non-, and negative responders differed statistically significant in terms of baseline performance of FICSIT-4 (FICSIT-4, modified SPPB), modified 30s CST

(modified 30s CST, modified SPPB), and modified SPBB (modified SPPB, see Table 28). The post-hoc analysis (see Table 28) revealed statistically significantly 1) worse performance of positive compared to negative responders for FICSIT-4 (FICSIT-4) and modified 30s CST (modified 30s CST); 2) worse performance of positive compared to non-responders for FICSIT-4 (modified SPPB), modified 30s CST (modified SPPB), and modified SPPB (modified SPPB).

Table 28. Differences in baseline physical and cognitive performance between positive, non-, and negative responders in the intervention group (per protocol analysis)

	Negative responders	**Non-responders**	**Positive responders**	**Between group difference**	**Post-hoc analysis**	
	Mean (SD)	Mean (SD)	Mean (SD)	$F(df_{numerator}, df_{denominator})$/ $Chi^2(df)$, p		
6m WT						
MMSE (n=86)	16.7 (4.1)	16.8 (4.3)	17.5 (4.9)	F(2,83)=0.252, p=0.777, η_p^2=0.006		
6m WT (n=86)	9.1 (2.9)	10.4 (3.4)	10.0 (3.1)	F(2,83)=1.277, p=0.284, η_p^2=0.030		
Modified SPPB (n=82)	7.5 (2.7)	6.4 (2.7)	6.7 (2.9)	F(2,86)=1.199, p=0.307, η_p^2=0.029		
TUG						
MMSE (n=85)	16.6 (3.8)	17.7 (4.5)	17.0 (5.0)	Chi^2(2)=1.061, p=0.588		
TUG (n=85)	19.8 (9.7)	19.3 (7.1)	24.1 (12.8)	Chi^2(2)=2.847, p=0.241		
Modified SPPB (n=81)	7.3 (3.0)	6.8 (2.2)	6.8 (3.0)	F(2,78)=0.305, p=0.738, η_p^2=0.008		
FICSIT-4						
MMSE (n=93)	17.3 (4.5)	15.8 (3.9)	17.1 (4.5)	F(2,90)=1.004, p=0.370, η_p^2=0.022		
FICSIT-4 (n=93)	3.0 (1.3)	2.4 (1.5)	1.3 (1.1)	Chi^2(2)=23.083, p<0.001	z=4.722, r=0.58 [a]	p<0.001,
					z=3.060, r=0.40 [b]	p=0.007,
Modified SPPB (n=88)	7.3 (2.9)	6.8 (3.0)	5.8 (2.5)	F(2,85)=2.447, p=0.093, η_p^2=0.054		

Modified 30s CST					
MMSE (n=69)	17.0 (3.8)	17.8 (4.4)	18.0 (5.0)	$Chi^2(2)=1.035$, p=0.596	
Modified 30s CST (n=69)	9.2 (3.9)	9.1 (2.6)	6.3 (2.8)	F(2,66)=5.244, p=0.008, $\eta_p^2=0.137$	p=0.011, MD=-2.81, CI_{95} [-5.08, -0.55] [a] p=0.028, MD=-2.81, CI_{95} [-5.36, -0.25] [b]
Modified SPPB (n=67)	7.5 (2.6)	7.6 (3.0)	5.9 (2.3)	F(2,64)=2.668, p=0.077, $\eta_p^2=0.077$	
Modified SPPB					
MMSE (n=76)	16.5 (3.8)	16.5 (4.3)	18.1 (4.7)	$Chi^2(2)=2.805$, p=0.246	
6m WT (n=76)	10.5 (4.8)	8.7 (2.3)	9.8 (3.0)	$Chi^2(2)=2.730$, p=0.255	
FICSIT-4 (n=76)	2.5 (1.3)	3.1 (1.3)	1.9 (1.5)	$Chi^2(2)=9.446$, p=0.009	z=3.051, p=0.007, r=0.43 [b]
Modified 30s CST (n=69)	8.0 (3.5)	9.6 (4.3)	6.8 (2.7)	F(2,66)=3.701, p=0.030, $\eta_p^2=0.101$	p=0.022, MD=-2.86, CI_{95} [-5.39, -0.34] [b]
Modified SPPB (n=76)	6.9 (2.9)	8.3 (2.9)	5.7 (2.3)	$Chi^2(2)=9.066$, p=0.011	z=2.980, p=0.003, r=0.42 [b]

6m WT: 6-meter walk test, CI_{95}: 95 % confidence interval, df: degrees of freedom, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, MMSE: Mini-Mental State Examination, modified 30s CST: 30-second chair stand test, modified SPPB: Short Physical Performance Battery, n: number of individuals with dementia, MD: mean difference, SD: standard deviation, TUG: Timed Up & Go Test

[a] post-hoc analysis: statistically significant worse performance of positive compared to negative responders

[b] post-hoc analysis: statistically significant worse performance of positive compared to non-responders

6.1.5 Discussion

The aim of the present study was to examine the impact of a MEP specially-designed for IWD by combining motor and cognitive tasks, on physical performance using a multicenter RCT study design. No statistically significant time*group effects were observed either in the intention-to-treat or the per protocol analysis. Thus, our hypothesis that physical performance of IWD changes after a 16-weeks MEP as compared to a control condition in which participants only received conventional treatment could not be confirmed.

This RCT had a strong focus on methodological correctness and may thus expand on previous literature which had some methodological limitations (Blankevoort et al.,

2010). In line with this, we applied two analysis procedures. However, both procedures have limitations regarding the evaluation of the outcomes. The intention-to-treat analysis might underestimate the real effects, whereas the per protocol analysis might overestimate these. For our samples, both analyses showed no effects of MEP on physical performance when considering the mean changes. One likely reason for the lack of overall effects might be the relatively high heterogeneity of our sample. High standard deviations were seen for all assessments in both IG and CG and might be a consequence of the high age and the associated different experiences of life as well as differences in disease stage and other comorbidities and constraints relating to the three basic abilities, i.e. mobility, balance, and strength. The high heterogeneity of samples of IWD in clinical trials was reported as limitation before (Cohen-Mansfield, 2000).

No effects in overall results of physical performance were reported in other RCT (Netz et al., 2007; Pitkälä, Pöysti et al., 2013). Lamb et al. (2018) observed an improved physical fitness after an aerobic and strength exercise training program of moderate to high intensity, but did not report any noticeable improvements in other clinical outcomes. Even though some literature reviews concluded that there is a trend for positive effects of a physical activity intervention on physical performance in IWD, the results of original research studies are still conflicting. In line with this, our study could also not establish a beneficial effect of the MEP on motor performance. There are several challenges with regard to methodological diversities especially in terms of interventions and assessments that need to be overcome by future studies. For example, physical activity interventions have several degrees of freedom like motor skills (endurance, strength etc.) and training parameters (intensity, duration etc.) and it remains unclear which motor skills or training parameters have the highest impact on fitness. One approach to solving this research gap could be to identify these relations before applying multimodal physical interventions. Furthermore, assessing motor skills is challenging in IWD. The assessment might be biased by cognitive impairments which may in turn lead to lower reliability (Blankevoort et al., 2013). Additionally, present assessments often do not measure the actual performance but are rather biased by mood or motivation of IWD. Thus, the claim for appropriate assessment tools that are specifically developed for IWD is meaningful and justified. In healthy older adults, there is much stronger evidence on the beneficial impact of physical activity interventions on physical

performance. Thus, there might be other factors that may be underlying this relationship and that have not yet been sufficiently investigated in IWD.

Conducting the responder/non-responder analysis in which we identified those IWD that respond to the intervention, revealed important insights that may be clinically meaningful. In fact, between 56 % of participants (for modified 30s CST) and 66 % of participants (for 6m WT and modified SPPB) experienced no change or even an improvement of their physical performance. These findings reflect the positive, subjective feedback we received from several participants, relatives, or employees of care facilities. Even maintaining the level of physical performance over the 16-weeks study period may be a sign for a positive impact of the MEP as rapid decline of motor performance is usually seen in dementia patients. Furthermore, findings in relation to the increase of SPPB values is of great clinical relevance as this assessment is related to mobility disability and may be a strong predictor of falls and death risk (Pavasini et al., 2016; Vasunilashorn et al., 2009; Veronese et al., 2014).

In addition, our results indicate that the benefit of the MEP varies depending on the level of physical performance at baseline. Those IWD that showed low performance in balance (FICSIT-4) and strength and function of lower extremities (modified 30s CST and modified SPPB) have a higher chance of experiencing positive changes in the same variables after the MEP. This is clinically relevant, and one may conclude that physical activity treatment should particularly be recommended to IWD with poor physical performance. On the other hand, it is also possible that the intensity of our exercise sessions was too low to have a measurable impact on those IWD that had a higher level of physical performance at baseline. We cannot rule out this assumption as the intervention was executed within groups. Even though we deliberately had small group sizes, exercise instructors may have referred to participants with low physical fitness when determining the intensity of the MEP sessions. This was important in order to ensure safety for all participants during the MEP, including those with lower physical fitness. However, the intensity may have been too low for participants with higher fitness level to experience any improvement in physical performance. The key to success may thus be a higher degree of individualization within the intervention.

There is a general assumption that lower levels of motor function are more pronounced in individuals with cognitive impairments as compared to cognitively unimpaired persons (Camicioli, Howieson, Oken, Sexton, & Kaye, 1998). However, in our sample, we

did not observe that poor motor function is associated with more severe cognitive impairment. Even though motor functions are affected by several other factors. The responders in balance as well as strength and function of lower extremities that have poor levels within these variables at baseline did not differ in the global cognition (MMSE). Our results cannot be compared to related work, because a responder analysis has not been conducted before.

The strengths of this study are the high-quality methodological approach as well as the precise documentation and reporting. Of note, our RCT had a large sample size as well as a MEP which proved to be a feasible and sustainable exercise program tailored to the characteristics and demands of IWD. This intervention was conducted across 36 care facilities, and most of these facilities implemented the MEP into their daily routine after the end of the study.

Despite of the methodological and planning efforts, some limitations pertain to our research. First, even though the MEP was carefully developed based on theoretical considerations, as well as the results of a pilot study (Thurm et al., 2011), and a literature review (Scharpf et al., 2013), it may not have allowed to sufficiently take into account the baseline level of participants. As the MEP was delivered in a group setting, some participants may not have reached the intensity threshold needed to induce any changes in motor fitness. Furthermore, some of the effects we observed could be due to the group setting, i.e. enhanced social interaction, or additional attention that participants received from the exercise instructors, rather than due to the MEP. This bias could be addressed in future studies by additional non-exercise groups of social interventions. Another limitation pertains to the differences concerning the sample sizes between intention-to-treat and per protocol analysis, which were relatively high as several IWD did not complete the assessment battery. Reasons for missing data in the assessments were severe disease, impaired walking ability, or refusal due to appointment scheduling conflict. It remained unclear if this refusal was usually due to excessive demands of the MEP, reluctance or lacking motivation or other daily conditions. The major limitation is that assessments to identify motor function that are frequently used in research studies are often not specifically designed for IWD. Designing tests to assess motor performance among IWD is critically important and should be considered in future studies among IWD.

6.1.6 Conclusion

This multicenter RCT aimed to identify the effects of a physical activity intervention on physical performance in IWD. Overall, there were no significant effects of the standardized MEP on motor performance. While keeping in mind the limitations of responder analysis, this analysis resulted in a considerable proportion of participants that responded to the MEP by maintaining or even improving their motor function. The main recommendation that can be derived from this RCT is the need to individualize physical activity interventions among IWD. This recommendation is supported by the following two aspects. First, lack of an impact of the MEP on physical performance along with the examination of characteristics of the sample is indicative of a high heterogeneity of participants in our sample. This heterogeneity was obvious in all motor functions and might result from the high age of study participants and the associated different experiences of life as well as differences in disease stage and other comorbidities and constraints. For this reason, we recommend to better adapt interventions by considering the individual needs when planning an appropriate exercise training. Second, findings from responder analysis showed that 56-66 % of participants showed either no decline or even an increase in physical performance. Particularly individuals with low baseline performance in balance as well as strength and function of lower extremities, seemed to benefit from the MEP and improve their performance in these tests. Thus, future research should focus on including the individual baseline performance in the planning of the training. There might be additional moderating or mediating factors that should be considered when planning an activity program. It can thus be beneficial to conduct other exploratory statistical procedures like cluster analyses to identify the most promising combination of individual characteristics. In conclusion, this RCT contributes to the body of research on the impact of physical activity interventions in IWD, and its results may inform the design and conduct of future intervention studies. Furthermore, this research serves as a starting point for our team’s future study which will examine the effects of an individualized physical activity program in IWD.

Abbreviations

6m WT: 6-meter walk test; ANOVA: analysis of variance; BMI: Body Mass Index; CG: control group; CIRS: Cumulative Illness Rating Scale; CONSORT: Consolidated Standards of Reporting Trials; FIC-SIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4; IG: intervention group; IWD: individuals with dementia; MEP: multimodal exercise program; MMSE: Mini-Mental State Examination; modified 30s CST: modified 30-second chair stand test; modified SPPB: modified Short Physical Performance Battery; RCT: randomized controlled trial/s; SPPB: Short Physical Performance Battery; TUG: Timed Up & Go Test

Declarations

Conflicts of Interest and Source of Funding

The authors declare that they have no conflicts of interests.

This project is financially supported by the Dietmar Hopp Foundation (St. Leon-Rot, Germany). The sponsor did not have any role in the design of the study, neither in its execution, the collection, analysis or interpretation of data, the decision to submit results nor in writing the report.

Acknowledgments

We are grateful to all participating in the study and to our research stuff. We would like to thank Luisa Appelles for the coordination of fieldwork and for entering data.

References

All references of manuscript IV are included in the List of References at the end of this thesis.

6.2 Effectiveness of a dementia-specific multimodal exercise program on gait performance in individuals with dementia

Manuscript V

Summary: Gait impairments have a high prevalence in IWD and are associated with an increased risk of falls (Allali & Verghese, 2017). Previous studies investigating the effectiveness of physical activity on gait performance in IWD show promising but inconsistent results. Moreover, research on impacts, e.g. characteristics of responders or underlying changes in motor and cognitive performance, is rare. Accordingly, there is a need for further high-quality studies. Related findings may also contribute to improving physical activity interventions. Manuscript V aims to investigate the effectiveness of physical activity on gait performance in IWD and to determine impacts on changes in gait performance. Herein, it focuses on research questions B2, B3, and B4, as well as on hypotheses 2-7.

Version: This is the author's original before peer review.

Publication status: submitted

Reference:

Trautwein, S., Barisch-Fritz, B., Scharpf, A., Ringhof, S., Stein, T., Krell-Roesch, J., & Woll, A. (submitted). Effects of a 16-week multimodal exercise program on gait performance in individuals with dementia: A multicenter randomized controlled trial.

Effects of a 16-Week Multimodal Exercise Program on Gait Performance in Individuals with Dementia: A Multicenter Randomized Controlled Trial[13]

Sandra Trautwein [a], Bettina Barisch-Fritz [a], Andrea Scharpf [a], Steffen Ringhof [b], Thorsten Stein [a], Janina Krell-Roesch [a, c], and Alexander Woll [a]

[a] Karlsruhe Institute of Technology, Institute of Sports and Sports Science, Karlsruhe, Germany

[b] University of Freiburg, Department of Sport and Sport Science, Freiburg, Germany

[c] Mayo Clinic, Translational Neuroscience and Aging Lab, Scottsdale, Arizona, USA

6.2.1 Abstract

Background: There is a high prevalence of gait impairments in individuals with dementia (IWD). Gait impairments are associated with increased risk of falls, disability, and economic burden. Only few studies have investigated the effectiveness of physical activity on gait performance in IWD, reporting promising but inconsistent results.

Objective: To investigate the effectiveness of a multimodal exercise program (MEP) on gait performance in IWD.

Methods: In this randomized controlled trial, we enrolled 319 IWD of mild to moderate severity. The intervention group participated in a 16-week MEP specifically tailored to IWD. We examined the effects of the MEP on spatiotemporal gait parameters and dual task costs by using GAITRite. Additionally, we compared characteristics between positive, non-, and negative responders, and investigated the impact of changes in underlying motor and cognitive performance.

Results: There were no statistically significant time*group effects on either spatiotemporal gait parameters or dual task costs. Differences in baseline gait performance, mobility, lower limb strength, and severity of cognitive impairments were observed between positive, non-, and negative responders. Changes in lower limb strength and function, mobility, executive function, attention, and working memory explained up to 39.4 % of the variance of changes in gait performance.

Conclusion: The effectiveness of a standardized MEP on gait performance in IWD was limited, probably due to the large heterogeneity of the sample. However, additional analyses revealed prerequisites of individual characteristics and impacts of changes

[13] Some minor formal adaptions were made to the version of manuscript V presented in this thesis to ensure uniform formatting.

in underlying motor and cognitive performance. Considering such factors may improve the effectiveness of a physical activity intervention among IWD.

Trial registration: DRKS00010538 (German Clinical Trial Register, date of registration: 01 June 2016, retrospectively registered).

Keywords: Exercise, dementia, gait analysis, physical functional performance, cognition

6.2.2 Introduction

Gait impairments represent a major public health concern (Valkanova & Ebmeier, 2017). Their prevalence increases with age, and more than 32 % of individuals aged 60 years and above have gait impairments (Mahlknecht et al., 2013) such as decreased walking speed, shortened stride length, and enhanced double support phase (Mc Ardle et al., 2017; Valkanova & Ebmeier, 2017; van Iersel et al., 2004). Compared to cognitively unimpaired older individuals, gait impairments are more prevalent in individuals with dementia (IWD), with an estimated 50 % of IWD being affected (Allali & Verghese, 2017; Allan et al., 2005). Motor impairments, such as reduced strength and postural control, may contribute to this increased prevalence of gait impairments in IWD (Alexander & Goldberg, 2005; Allali & Verghese, 2017). Moreover, gait is not merely an automated motor activity but requires input from the cerebellum, the motor cortex, and the basal ganglia, as well as an intact sensory feedback (Amboni et al., 2013; Hausdorff, 2007; Valkanova & Ebmeier, 2017). Thus, dementia-related pathological changes in these brain structures may also contribute to gait impairments (Mc Ardle et al., 2017).

Both gait and cognitive impairments are associated with an increased risk of falls (Montero-Odasso et al., 2012). Accordingly, the incidence of falls in IWD is two to three times higher than in cognitively unimpaired older individuals (Allali & Verghese, 2017; Montero-Odasso et al., 2012; Valkanova & Ebmeier, 2017). Furthermore, the various health-related and economic consequences of falls, such as higher rates of institutionalization, disability, morbidity, mortality, and increased financial burden (Allali & Verghese, 2017; Valkanova & Ebmeier, 2017), underline the need of interventions focusing on improving or maintaining gait performance in IWD. Indeed, various pharmacological and non-pharmacological interventions to improve gait performance and reduce falls in older adults have been studied.

Physical activity interventions have shown to be effective in cognitively unimpaired older individuals and may also be beneficial for IWD (Allali & Verghese, 2017). However, to date, only few studies have evaluated the effectiveness of physical activity on gait performance in IWD. These studies show promising but inconsistent results. For example, seven studies observed positive effects of physical activity on walking speed as assessed through short distance walk tests (Ahn & Kim, 2015; Aman & Thomas, 2009; Bossers, Scherder et al., 2014; Hauer et al., 2012; Manckoundia et al., 2014; Rolland et al., 2007; Toulotte et al., 2003), whereas fifteen studies did not report statistically significant findings (Bossers et al., 2015; Cadore et al., 2014; Dawson et al., 2019; Hageman & Thomas, 2002; Hauer et al., 2017; Junge et al., 2018; Kuiack et al., 2004; McCaffrey et al., 2014; Sobol et al., 2016; Souto Barreto et al., 2017; Steinberg et al., 2009; Tay et al., 2016; Telenius et al., 2015a; Thomas & Hageman, 2003; Toots et al., 2017). Furthermore, ten studies applied an instrumented gait analysis, and mainly reported positive effects of physical activity on stride length (Coelho et al., 2013; Kemoun et al., 2010; J.-S. Kim et al., 2017; Orcioli-Silva et al., 2018; Pedrinolla et al., 2018; Perrochon et al., 2015; Schwenk, Dutzi et al., 2014; Schwenk, Zieschang et al., 2014), stride time (Orcioli-Silva et al., 2018; Schwenk, Zieschang et al., 2014), step time (J.-S. Kim et al., 2017), double support time (Kemoun et al., 2010), and stride frequency (Perrochon et al., 2015). In contrast, no effects were found on step length (Pedrinolla et al., 2018; Suttanon et al., 2013), step width (Schwenk, Zieschang et al., 2014; Suttanon et al., 2013), and percent of single support (Pedrinolla et al., 2018). Inconsistent results exist for walking speed (Kemoun et al., 2010; J.-S. Kim et al., 2017; Pedrinolla et al., 2018; Perrochon et al., 2015; Ries et al., 2015; Schwenk, Dutzi et al., 2014; Schwenk, Zieschang et al., 2014; Suttanon et al., 2013), stride speed (Coelho et al., 2013; Orcioli-Silva et al., 2018), percent of double support (Orcioli-Silva et al., 2018; Pedrinolla et al., 2018; Schwenk, Zieschang et al., 2014), and cadence (Coelho et al., 2013; Hauer et al., 2012; J.-S. Kim et al., 2017; Orcioli-Silva et al., 2018; Schwenk, Dutzi et al., 2014; Schwenk, Zieschang et al., 2014). Findings of studies investigating dual task conditions are also inconsistent and do not allow meaningful conclusions (Cadore et al., 2014; Coelho et al., 2013; Junge et al., 2018; Orcioli-Silva et al., 2018; Sobol et al., 2016; Tay et al., 2016). Thus, more research is needed to better understand the potentially beneficial effects of physical activity on gait performance in IWD in both single and dual task conditions.

Most previous studies conducted multimodal physical activity interventions, including strength, balance, and aerobic exercises (Aman & Thomas, 2009; Bossers et al., 2015; Bossers, Scherder et al., 2014; Cadore et al., 2014; Coelho et al., 2013; Dawson et al., 2019; Hauer et al., 2012; Hauer et al., 2017; Junge et al., 2018; Kemoun et al., 2010; Manckoundia et al., 2014; Orcioli-Silva et al., 2018; Pedrinolla et al., 2018; Perrochon et al., 2015; Rolland et al., 2007; Schwenk, Dutzi et al., 2014; Schwenk, Zieschang et al., 2014; Souto Barreto et al., 2017; Steinberg et al., 2009; Suttanon et al., 2013; Tay et al., 2016; Telenius et al., 2015a; Toots et al., 2017; Toulotte et al., 2003). Given the relationship between motor, cognitive and gait performance, as well as the positive impacts of cognitive training programs (Amboni et al., 2013; Valkanova & Ebmeier, 2017), interventions combining physical and cognitive activity, may be most promising for improving gait performance in IWD (J. A. Cohen et al., 2016). Indeed, studies combining physical and cognitive activity predominantly reported beneficial effects on gait performance (Cadore et al., 2014; Coelho et al., 2013; Orcioli-Silva et al., 2018; Schwenk, Zieschang et al., 2014). However, these studies had no randomized controlled trial designs (Cadore et al., 2014; Coelho et al., 2013; Orcioli-Silva et al., 2018), did not use instrumented gait analysis systems (Cadore et al., 2014), or focused on dual task exercises while not considering other cognitive tasks (Schwenk, Zieschang et al., 2014). This research gap emphasizes the need for additional investigations.

When aiming to improve the effectiveness of physical activity interventions on gait performance in IWD, it is also important to consider and identify determinants that may potentially impact the association between physical activity and subsequent changes in gait performance. However, research on such prerequisites, e.g. specific characteristics of participants that may determine which participants are most likely to benefit from specific physical activity interventions, is rare. With regard to the expected direct and indirect effects of physical activity (see Figure 9), little is known as to how intervention-induced changes in underlying motor and cognitive performance may be related to changes in gait performance in IWD. As both motor and cognitive impairments explain the increased prevalence of gait impairments in IWD (Allali & Verghese, 2017), potential impacts of both factors are possible. Based on theoretical considerations, associations between changes in gait performance with changes in balance, mobility, strength and function of lower limbs (Alexander & Goldberg, 2005) as well as with

changes in executive function, attention, and working memory (Sheridan & Hausdorff, 2007; Yogev-Seligmann et al., 2008) can be assumed.

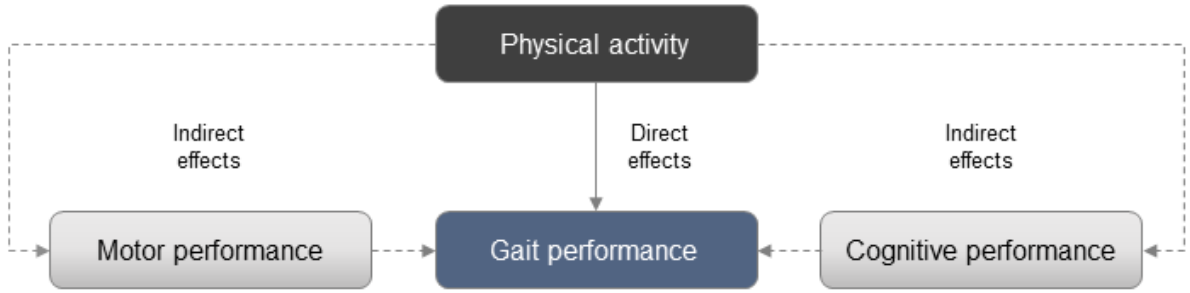


Figure 9. Direct and indirect effects of physical activity on gait performance.

Therefore, the primary aim of the current study was to determine the effectiveness of a multimodal exercise program (MEP), which combined both motor and cognitive tasks, on gait performance and dual task costs in IWD residing in care facilities. We hypothesized that a 16-week MEP, in addition to conventional treatment, is more effective in reducing the decline in gait performance in IWD than conventional treatment alone. Our secondary aim was to identify determinants that may affect the effectiveness of the MEP, by examining differences in characteristics closely related to gait performance between positive, non-, and negative responders. Furthermore, we also investigated impacts of intervention-induced changes in underlying motor and cognitive performance on changes in gait performance.

6.2.3 Methods

For this manuscript, we followed the guidelines and recommendations of the Consolidated Standards of Reporting Trials statements (Moher et al., 2010; Schulz et al., 2010). The reader is referred to the published study protocol for a detailed description of study design and methods (Trautwein et al., 2017). The following sections will only provide a brief summary of study methods. Further information is available in the German National Register of Clinical Trials (DRKS00010538), where we retrospectively registered this study. The ethics committee of the Karlsruhe Institute of Technology (Karlsruhe, Germany) granted ethical approval.

6.2.3.1 Study Design

We performed a multicenter randomized controlled trial with baseline and post-intervention assessments. We allocated participants to an intervention group (IG) or control group (CG) with an allocation ratio of 2:1 using minimization software (MinimPy, Version 0.3; Saghaei & Saghaei, 2011). If possible, we blinded investigators to group allocation. However, it was not possible to blind participants.

6.2.3.2 Participants

A power analysis (G*Power 3, Version 3.1.9.2 [Faul et al., 2007], two-factor analysis of variance [ANOVA] with repeated measurements, two groups, two measurements, $\alpha=0.05$, $1-\beta=0.80$, $\eta^2=0.01$) revealed a required total sample size of 200 participants. Considering various potential reasons for dropout, missing data, and low adherence to the MEP, we set the sample size to 405 participants. Participants were recruited from 36 care facilities in South-Western Germany. Employees of the respective care facilities identified eligible participants, which had to fulfill the following inclusion and exclusion criteria:

Inclusion criteria: a) diagnosis of dementia or “suspected” dementia; b) Alzheimer’s disease, vascular dementia, or other primary dementia; c) mild to moderate severity of dementia (Mini-Mental State Examination [MMSE] score: 10-24); d) age above 65 years; e) walking ability of about ten meters with or without walking aid; and f) clearance from general practitioner.

Exclusion criteria: a) secondary dementia; b) other severe cognitive impairments; c) other severe neurological diseases; d) any severe acute diseases; and e) severe motor impairments.

Based on these inclusion and exclusion criteria, we verified the eligibility of participants at baseline assessment. Furthermore, we obtained written informed consent prior to the study from all participants or their legal guardians, respectively. Participation in this study was voluntary.

6.2.3.3 Sample characteristics

The employees of the care facilities documented characteristics of participants including sex, year of birth, diagnosis of dementia, etiology of dementia, walking aids, depression, Cumulative Illness Rating Scale (CIRS) (Linn et al., 1968), and medication intake within two weeks of baseline assessments. Whenever possible, we asked physicians to retrospectively provide any missing information on their patients. In addition, we measured body mass and height in all participants.

6.2.3.4 Intervention

Participants in the IG underwent an MEP combing motor (i.e. strength, balance, endurance, and flexibility) and cognitive tasks (i.e. memory, attention, language, and executive function). The MEP was tailored to fit the specific needs and characteristics of IWD, and was delivered in the care facilities by instructors who had been specifically trained for the purpose of this study. In order to provide a sense of security for participants, the MEP included a ritualization that ensured an identical sequence for all sessions. It was implemented by asking participants to imagine a journey while performing appropriate motor and cognitive exercises. The MEP took place twice a week over a period of 16 weeks. Sessions had a duration of 60 minutes including 45 minutes of physical exercise. The MEP was delivered in a group setting and was mainly performed in a seated position with medium to submaximal intensity. During the course of the 16 weeks, we increased the intensity of the sessions as well as the degree of motor and cognitive requirements. Both CG and IG participants received individually tailored conventional treatment (e.g. medication, care, or therapeutic applications) as part of standard care in their care facilities.

6.2.3.5 Outcome measurements

We examined gait performance as outcomes of interest with various spatiotemporal gait parameters of the right leg: walking speed (m/sec), stride length (cm), stride time (sec), double support phase (% of stride time), and stance phase (% of stride time). Gait analysis was performed using the electronic gait analysis system GAITRite (CIR Systems Inc., Franklin, USA, active length of 4.88 meter), which has been shown to be reliable in IWD (Ries et al., 2009; Wittwer et al., 2008). All participants underwent gait analysis in single and two dual task conditions (i.e. counting backwards starting

from 50 and naming animals while walking) to also assess dual tasks costs of walking while talking.

To eliminate acceleration and deceleration during the recording, we asked participants to start walking two meters in front of the GAITRite system and to stop walking two meters behind the system (Kressig & Beauchet, 2006). While walking at comfortable speed, participants were allowed to use walking aids as applied in everyday life. Instructions were repeated if necessary. We asked participants to repeat all conditions up to five times to generate three valid walks. Valid trials consisted of a minimum of six consecutive steps of steady-state walking, and complied with satisfactory cognitive performance in dual task conditions. For statistical analysis, we considered the trial with the smallest difference to mean walking speed of all valid trials of one condition. We calculated dual-task costs using the equation $\frac{dual\ task - single\ task}{single\ task} \cdot 100$ (Abernethy, 1988; Schwenk et al., 2010).

In order to analyze differences between positive, negative, and non-responders, as well as impacts of changes in underlying motor and cognitive performance on changes in gait performance, we determined related outcomes using the motor and cognitive assessments displayed in Table 29. The reader is referred to the published study protocol (Trautwein et al., 2017) for a detailed description of all assessments.

Table 29. Motor and cognitive assessments to analyze differences between positive, non-, and negative responders, as well as impacts of changes in underlying motor and cognitive performance on changes in gait performance

Outcome	Assessment
Balance	Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4 [score] (Rossiter-Fornoff et al., 1995)
Mobility	Timed Up & Go Test (time in s) (Podsiadlo & Richardson, 1991)
Strength and function of lower limb	Modified 30-Second Chair-Stand Test [number of repetitions] (Blankevoort et al., 2013; Jones et al., 1999) Modified Short Physical Performance Battery [score] (Guralnik, Simonsick et al., 1994)*
Global cognition	Mini-Mental State Examination [score] (Folstein et al., 1975)
Executive function and visual-spatial cognition	Clock Drawing Test [adapted Sunderland score] (Mendes-Santos, Mograbi, Spenciere, & Charchat-Fichman, 2015; Shulman et al., 1986)
Executive function and processing speed	Trail Making Test A [established score considering time, final number, and non-corrected mistakes, a higher score indicates better performance] (Reitan, 1958, 1992)
Attention and working memory	Digit Span forward and backward [number of correct digit spans] (Wilde et al., 2004)

* standing balance, gait speed, and modified 5 times sit-to-stand (with using arms)

6.2.3.6 Statistical analysis

Statistical analysis was performed using IBM SPSS Version 25 (IBM Corporation, Armonk, USA). We ran a per protocol analysis including participants who had a MEP adherence of at least 75 % (only in IG) and a complete assessment of spatiotemporal gait parameters in at least one condition. Additionally, we implemented an intention-to-treat analysis and used multiple imputation (fully conditional specification imputation method, ten imputations, and ten iterations) to account for missing data. However, we did not impute data of deceased participants. To ensure plausibility of imputed data, we defined the following constraints: gait performance as both outcome and predictor variable, adherence as well as related motor and cognitive performance as predictor variables only; minimum and maximum values according to observed range in each variable; rounding according to original data; and 100 maximal case draws, ten maximal parameter draws. We considered pooled results as provided by SPSS or reported ranges observed throughout the imputations, if SPSS did not support the pooling procedure, as final point estimates.

Required assumptions were tested before performing statistical analyses. For comparison of baseline values and sample characteristics between IG and CG, we used Chi-square tests, Mann-Whitney-U-Tests, and unpaired T-Tests according to the scaling of the investigated outcome. We analyzed treatment effects using two-factor ANOVA with repeated measurements (time*group effects), and supplemented paired T-Tests (within group time effects). A two-sided p-value less than 0.05 indicated statistical significance. To account for multiple comparisons, we also considered adjusted significance levels using Bonferroni-Holm correction in primary analyses. Additionally, we calculated 95 % confidence intervals of differences between baseline and post-intervention assessments and partial Eta^2.

In secondary exploratory analyses, we included walking speed, stride length, and double support of the per protocol IG sample and determined differences in baseline performance (i.e. balance, mobility, strength and function of lower limbs, executive function, attention, and working memory) and selected sample characteristics (i.e. severity of cognitive impairments, etiology of dementia, and use of walking aids) between positive, non-, and negative responders using Chi-square tests, Kruskal-Wallis-Tests, and one-factor ANOVA. For post-hoc analyses, we used Dunn-Bonferroni-Tests and Tukey-Kramer post-hoc tests, respectively. R and partial Eta^2 served as effect sizes.

We defined positive responders as those participants, who improved their gait performance at least 10 % during the 16-week MEP, while negative responders showed a decline of at least 10 % in gait performance, and non-responders were participants with less than 10 % improvement or decline. This definition was based on percentage minimal detectable changes at 95 % confidence interval of considered spatiotemporal gait parameters which ranged between 7 % and 12 % in a reliability study using GAITRite (Trautwein, Maurus et al., 2019; Wittwer et al., 2008). The minimal detectable change is a measure of absolute reliability, which delineates “expected” from “true” changes in performance (Ries et al., 2009). Moreover, we assessed the potential impact of changes in underlying motor and cognitive performance on changes in gait performance using multiple linear regression models with stepwise selection. Based on theoretical assumptions, we considered changes in balance, mobility, strength and function of lower limbs, executive function, attention, and working memory as independent variables. The calculated effect size is f^2.

6.2.4 Results

6.2.4.1 Recruitment and flow of participants

Recruitment took place between March 2015 and March 2017. We screened 600 IWD for eligibility, of whom 319 were enrolled in the study. 201 participants were allocated to the IG and 118 to the CG. There was a dropout rate of 8 % in both IG and CG, respectively. There were no statistically significant differences in sample characteristics or baseline measurements between participants who dropped out versus those who completed the study. The mean adherence in the IG was 62 %. 107 participants (53 %) of the IG completed the MEP in accordance with the study protocol, i.e. defined by a minimum adherence of at least 75 % of all sessions. 65 % of participants in the IG and 62 % of participants in the CG completed at least one condition of gait analysis at baseline and post-intervention assessment. Based on the above-mentioned criteria, 163 participants could be considered for the per protocol analysis. Even though we extended our initially planned recruitment phase for an additional year, we were not successful in reaching our intended sample size of 405 participants. This is due to the fact that the number of participants who did not fulfill our inclusion and exclusion criteria was much larger than expected. Nevertheless, a sensitivity analysis using G*Power 3 (Version 3.1.9.2; Faul et al., 2007) showed that we were still able to detect small to

medium effects with our actual sample size (η^2=0.013 to 0.018). Figure 10 displays the flow of participants and states the reasons for dropouts and non- participations in assessments.

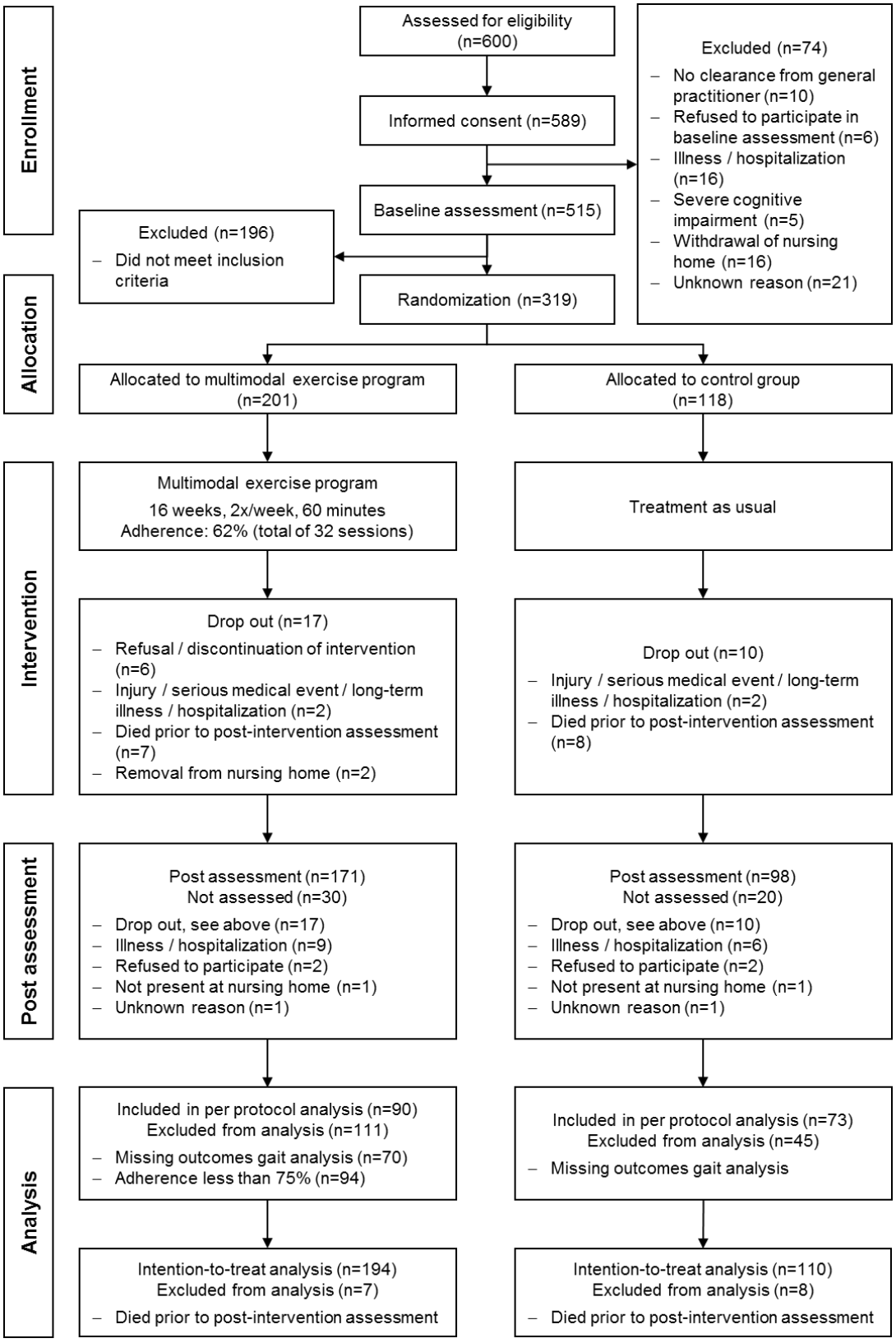


Figure 10. Flow of participants (n: number).

6.2.4.2 Sample characteristics

Table 30 provides an overview of the characteristics of participants at baseline (per protocol analysis; see Supplementary Table 1 for sample characteristics of the intention-to-treat analysis). The participants' mean (SD) age was 86 (6) years and 84 % of participants were female. A mean (SD) MMSE score of 17 (4) (range: 10-24) indicated a mild to moderate severity of dementia at baseline assessment. 77 % of participants were dependent on walking aids, and a mean (SD) CIRS morbidity index of 9 (4) as well as an average (SD) of 7 (4) required medications may indicate presence of medical comorbidities in the sample. We observed no statistically significant differences in characteristics between the IG and CG, except for the number of medications and intake of antidepressants.

Table 30. Sample characteristics of participants at baseline (per protocol analysis)

	Total sample [n=163]	**Intervention group** [n=90]	**Control group** [n=73]	**Group differences** [t(df)/z/Chi²(df), p]
Age, years [M (SD), range]	85 (6), 67-98	85 (7), 67-97	86 (5), 70-98	t(160.931)=1.918, p=0.057
Sex, female	84 %	82 %	86 %	Chi²(1)=0.500, p=0.479
Diagnosis of dementia				Chi²(2)=1.944, p=0.378
- yes	72 %	71 %	73 %	
- no	16 %	19 %	12 %	
- unknown	12 %	10 %	15 %	
Type of dementia				Chi²=5.693, p=0.199
- Alzheimer's disease	15 %	14 %	15 %	
- Vascular dementia	17 %	21 %	11 %	
- Mixed dementia	3 %	2 %	4 %	
- other	1 %	2 %	0 %	
- unknown	36 %	31 %	43 %	
- no/unknown diagnosis	28 %	29 %	27 %	
MMSE [M (SD), range]	17 (4), 10-24	17 (4), 10-24	17 (4), 10-24	t(160.446)=0.317, p=0.752
Use of walking aid				Chi²(2)=4.644, p=0.098
- walker	69 %	62 %	77 %	
- waking stick/s	9 %	9 %	8 %	
- no walking aid	23 %	29 %	15 %	
Depression				Chi²(2)=2.118, p=0.347
- yes	26 %	30 %	21 %	
- no	54 %	52 %	56 %	
- unknown	20 %	18 %	23 %	

CIRS [M (SD), range]				
- Morbidity Index	9 (4), 1-26	9 (4), 1-20	8 (5), 2-26	t(101)=-0.633, p=0.528
- Severity Index	1.5 (0.4), 1-3 not available for 37 %	1.5 (0.4), 1-3 not available for 31 %	1.5 (0.4), 1-3 not available for 44 %	z=-0.247, p=0.805
Number of medications [M (SD), range]	7 (4), 0-27 unknown in 21 %	8 (4), 1-27 unknown in 22 %	5 (3), 0-12 unknown in 21 %	**t(126)=-3.627, p<0.001**
Antidementives				Chi²(2)=5.473, p=0.065
- yes	23 %	29 %	16 %	
- no	44 %	37 %	53 %	
- unknown	33 %	34 %	30 %	
Antidepressants				**Chi²(2)=10.043, p=0.007**
- yes	25 %	34 %	14 %	
- no	37 %	30 %	47 %	
- unknown	37 %	36 %	40 %	
Height, cm [M (SD), range]	156.6 (8.3), 140.5-186.0 unknown in 0.6 %	156.9 (8.4), 142.5-186.0 unknown in 1 %	156.3 (8.2), 140.5-176.5	t(160)=-0.421, p=0.674
Weight, kg [M (SD), range]	68.3 (12.8), 41.3-125.0 unknown in 3 %	69.9 (12.8), 46.2-125.0 unknown in 3 %	66.3 (12.6), 41.3-99.4 unknown in 3 %	t(156)=-1.745, p=0.083
BMI, kg/m² [M (SD), range]	27.8 (4.5), 17.6-48.5 unknown in 3 %	28.4 (4.4), 19.7-48.5 unknown in 3 %	27.1 (4.6), 17.6-36.5 unknown in 3 %	t(156)=-1.801, p=0.074

BMI: Body Mass Index, CIRS: Cumulative Illness Rating Scale, df: degree of freedom, M: mean, MMSE: Mini-Mental State Examination, n: number, SD: standard deviation

Statistically significant results appear bold

6.2.4.3 Effects of the multimodal exercise program on spatiotemporal gait parameters

6.2.4.3.1 Per protocol analysis

Participants of the IG (per protocol sample) had a mean adherence of 92 %. Table 31 presents baseline and post-intervention values, differences between baseline and post-intervention assessments, group differences at baseline, within group time effects, and time*group effects including effect sizes of spatiotemporal gait parameters for single and dual task conditions as well as dual task costs. We did not observe any statistically significant time*group effects.

Table 31. Effects of the multimodal exercise program on spatiotemporal gait parameters and dual task costs (per protocol analysis)

		Baseline [M (SD)]	**Group differences at baseline** [t(df), p]	**Post** [M (SD)]	**Difference post – baseline** [M (SD), [CI_{95}]]	**Within group time effects** [t(df), p]	**Time*group effects** $F(df_{numerator}, df_{denominator})$, p	Effect size η_p^2
Single task (IG: n=89, CG: n=73)								
Walking speed, m/sec	IG	0.67 (0.19)	t(160)=-1.659, p=0.099	0.65 (0.22)	-0.02 (0.13), [-0.05, 0.00]	t(88)=1.787, p=0.077	F(1,160)=0.036, p=0.849	0.000
	CG	0.62 (0.19)		0.60 (0.20)	-0.02 (0.13), [-0.05, 0.01]	t(72)=1.373, p=0.174		
Stride length, cm	IG	82.6 (19.7)	t(159.875)=-0.842, p=0.401	80.5 (21.2)	-2.1 (10.9), [-4.4, 0.2]	t(88)=1.825, p=0.071	F(1,160)=0.030, p=0.863 [a]	0.000
	CG	80.2 (15.7)		77.8 (16.9)	-2.4 (10.4), [-4.8, 0.0]	t(72)=1.973, p=0.052		
Stride time, sec	IG	1.3 (0.2)	**t(131.361)=2.346, p=0.020**	1.3 (0.2)	0.0 (0.2), [0.0, 0.1]	t(88)=-1.571, p=0.120	F(1,160)=0.195, p=0.660 [a]	0.001
	CG	1.3 (0.2)		1.4 (0.3)	0.0 (0.2), [0.0, 0.1]	t(72)=-0.853, p=0.397		
Double support, % of stride time	IG	38.0 (8.1)	t(160)=1.289, p=0.199	39.2 (8.4)	1.1 (4.9), [0.1, 2.2]	**t(88)=-2.182, p=0.032**	F(1,160)=0.005, p=0.943	0.000
	CG	39.6 (7.4)		40.8 (7.3)	1.2 (4.9), [0.0, 2.3]	**t(72)=-2.070, p=0.042**		
Stance phase, % of stride time	IG	68.9 (4.1)	t(160)=1.368, p=0.173	69.5 (4.4)	0.6 (2.5), [0.1, 1.1]	**t(88)=-2.208, p=0.030**	F(1,160)=0.004, p=0.949	0.000
	CG	69.8 (4.1)		70.4 (4.2)	0.6 (3.1), [-0.2, 1.3]	t(72)=-1.543, p=0.127		
Dual task, counting backwards (IG: n=62, KG: n=52)								
Walking speed, m/sec	IG	0.55 (0.16)	**t(112)=-2.236, p=0.027**	0.54 (0.16)	-0.02 (0.14), [-0.06, 0.02]	t(61)=1.001, p=0.321	F(1,112)=0.101, p=0.752	0.001
	CG	0.48 (0.17)		0.47 (0.16)	-0.01 (0.15), [-0.05, 0.03]	t(51)=0.470, p=0.641		
Stride length, cm	IG	78.2 (19.1)	t(112)=-1.407, p=0.162	78.8 (19.5)	0.5 (11.6), [-2.4, 3.5]	t(61)=-0.359, p=0.721	F(1,112)=0.193, p=0.661	0.002
	CG	73.5 (16.6)		75.1 (16.6)	1.6 (15.3), [-2.6, 5.9]	t(51)=-0.773, p=0.443		
Stride time, sec	IG	1.5 (0.3)	**t(95.044)=2.446, p=0.016**	1.5 (0.4)	0.1 (0.4), [0.0, 0.2]	t(61)=-1.605, p=0.114	F(1,112)=0.253, p=0.616 [a, b]	0.002
	CG	1.6 (0.3)		1.7 (0.5)	0.1 (0.4), [0.0, 0.2]	**t(51)=-2.149, p=0.036**		
Double support, % of stride time	IG	40.6 (9.3)	**t(112)=2.110, p=0.037**	41.3 (8.7)	0.7 (5.9), [-0.8, 2.2]	t(61)=-0.998, p=0.322	F(1,112)=0.042, p=0.839	0.000
	CG	44.2 (9.1)		44.7 (9.0)	0.5 (7.4), [-1.6, 2.6]	t(51)=-0.481, p=0.632		

Stance phase, % of stride time	IG	70.3 (5.2)	t(112)=1.850, p=0.067	70.4 (4.7)	0.2 (2.9), [-0.6, 0.9]	t(61)=-0.414, p=0.680	F(1,112)=0.009, p=0.925 [b]	0.000
	CG	72.1 (5.2)		72.3 (5.2)	0.2 (4.3), [-1.0, 1.4]	t(51)=-0.359, p=0.721		
Dual-task costs, counting backwards (IG: n=62, KG: n=52)								
Walking speed, %	IG	-20.5 (15.2)	t(112)=-1.105, p=0.271	-21.0 (15.5)	-0.6 (16.6), [-4.8, 3.6]	t(61)=0.278, p=0.782	F(1,112)=0.053, p=0.818	0.000
	CG	-23.5 (14.0)		-23.4 (17.3)	0.1 (17.3), [-4.7, 5.0]	t(51)=-0.061, p=0.952		
Stride length, %	IG	-8.8 (11.7)	t(112)=-0.853, p=0.395	-7.0 (9.4)	1.7 (13.0), [-1.6, 5.0]	t(61)=-1.042, p=0.302	F(1,112)=0.759, p=0.386 [b]	0.007
	CG	-10.5 (9.9)		-6.6 (14.6)	3.9 (13.6), [0.1, 7.7]	**t(51)=-2.064, p=0.044**		
Stride time, %	IG	17.3 (17.3)	t(112)=0.806, p=0.422	21.0 (23.9)	3.7 (25.7), [-2.8, 10.2]	t(61)=-1.130, p=0.263	F(1,112)=0.257, p=0.613	0.002
	CG	19.9 (17.1)		25.9 (24.3)	6.0 (22.8), [-0.3, 12.4]	t(51)=-1.905, p=0.062		
Double support, %	IG	11.0 (14.3)	t(112)=0.305, p=0.761	11.9 (12.8)	1.0 (15.5), [-3.0, 4.9]	t(61)=-0.491, p=0.625	F(1,112)=0.081, p=0.776	0.001
	CG	11.7 (10.8)		11.8 (14.8)	0.1 (15.1), [-4.1, 4.3]	t(51)=-0.069, p=0.945		
Stance phase, %	IG	3.1 (4.5)	t(112)=0.095, p=0.924	2.9 (3.6)	-0.2 (4.8), [-1.4, 1.1]	t(61)=0.252, p=0.802	F(1,112)=0.130, p=0.719	0.001
	CG	3.1 (3.7)		3.3 (4.8)	0.2 (5.5), [-1.3, 1.7]	t(51)=-0.254, p=0.800		
Dual task, naming animals (IG: n=61, KG: n=59)								
Walking speed, m/sec	IG	0.45 (0.14)	t(118)=-1.797, p=0.075	0.43 (0.13)	-0.01 (0.12), [-0.04, 0.02]	t(60)=0.805, p=0.424	F(1,118)=0.972, p=0.326	0.008
	CG	0.40 (0.14)		0.41 (0.13)	0.01 (0.12), [-0.02, 0.04]	t(58)=-0.593, p=0.555		
Stride length, cm	IG	70.4 (18.1)	t(118)=-1.415, p=0.160	71.2 (17.7)	0.9 (11.0), [-2.0, 3.7]	t(60)=-0.620, p=0.538	F(1,118)=0.040, p=0.841	0.000
	CG	65.9 (16.3)		66.3 (14.9)	0.4 (13.2), [-3.0, 3.9]	t(58)=-0.252, p=0.802		
Stride time, sec	IG	1.6 (0.4)	t(118)=1.480, p=0.141	1.7 (0.4)	0.1 (0.3), [0.0, 0.2]	t(60)=-1.823, p=0.073	F(1,118)=3.448, p=0.066	0.028
	CG	1.7 (0.5)		1.7 (0.5)	0.0 (0.3), [-0.1, 0.1]	t(58)=0.801, p=0.426		
Double support, % of stride time	IG	45.9 (9.4)	t(118)=1.526, p=0.130	45.2 (8.6)	-0.7 (7.1), [-2.5, 1.1]	t(60)=0.758, p=0.452	F(1,118)=0.085, p=0.771	0.001
	CG	48.5 (9.5)		48.2 (8.6)	-0.3 (7.2), [-2.2, 1.6]	t(58)=0.326, p=0.746		
Stance phase, % of stride time	IG	72.4 (4.7)	**t(118)=2.233, p=0.027**	72.3 (4.4)	-0.1 (3.9), [-1.1, 0.9]	t(60)=0.241, p=0.810	F(1,118)=0.107, p=0.744	0.001
	CG	74.5 (5.5)		74.1 (5.1)	-0.3 (3.7), [-1.3, 0.6]	t(58)=0.727, p=0.470		

Dual-task costs, naming animals (IG: n=60, KG: n=59)								
Walking speed, %	IG	-34.4 (15.9)	t(117)=-0.520, p=0.604	-32.4 (18.6)	2.0 (19.7), [-3.1, 7.1]	t(59)=-0.776, p=0.441	F(1,117)=0.696, p=0.406 [a]	0.006
	CG	-35.9 (16.0)		-31.2 (14.6)	4.8 (16.6), [0.4, 9.1]	**t(58)=-2.204, p=0.032**		
Stride length, %	IG	-17.9 (11.1)	t(117)=-0.406, p=0.685	-14.3 (12.4)	3.7 (15.4), [-0.3, 7.6]	t(59)=-1.842, p=0.070	F(1,117)=0.007, p=0.931	0.000
	CG	-18.8 (12.5)		-14.9 (12.7)	3.9 (12.9), [0.5, 7.3]	**t(58)=-2.316, p=0.024**		
Stride time, %	IG	30.1 (27.1)	t(117)=0.326, p=0.745	33.2 (29.8)	3.2 (30.3), [-4.7, 11.0]	t(59)=-0.807, p=0.423	F(1,117)=2.558, p=0.112 [b]	0.021
	CG	31.6 (24.6)		27.2 (22.7)	-4.4 (20.4), [-9.7, 0.9]	t(58)=1.669, p=0.100		
Double support, %	IG	25.4 (17.6)	t(117)=-0.643, p=0.522	21.5 (17.8)	-3.9 (21.7), [-9.5, 1.7]	t(59)=1.401, p=0.166	F(1,117)=0.005, p=0.946	0.000
	CG	23.4 (15.6)		19.3 (16.2)	-4.2 (15.7), [-8.3, -0.1]	**t(58)=2.027, p=0.047**		
Stance phase, %	IG	6.3 (4.5)	t(117)=0.832, p=0.407	5.5 (5.0)	-0.7 (6.0), [-2.3, 0.8]	t(59)=0.986, p=0.337	F(1,117)=0.389, p=0.534	0.003
	CG	7.0 (4.9)		5.6 (4.6)	-1.4 (4.8), [-2.6, -0.1]	**t(58)=2.183, p=0.033**		

CG: control group, CI_{95}: 95 % confidence interval, df: degrees of freedom, IG: intervention group, M: mean, n: number, SD: standard deviation

[a] variance homogeneity not fulfilled, [b] covariance homogeneity not fulfilled

Statistically significant results appear bold for α=0.05. When considering adjusted significance levels using Bonferroni-Holm correction for multiple comparisons, no statistically significant results were observed.

6.2.4.3.2 Intention-to-treat analysis

Missing data analysis showed an amount of missing data ranging between 8.5 % (single task condition at baseline) and 47.6 % (dual task counting backwards at post-intervention assessment). With respect to gait performance, 194 of 319 records were incomplete. Reasons for missing values included not participating at post-intervention assessment (see Figure 10), weak physical condition, medical constrains, refusal, discontinuation of the assessment, invalid gait or dual task performance, and technical problems. Participants with incomplete data showed lower cognitive, motor, and gait performance, were older, required more medication, and had worse CIRS scores, depending on walking condition and time of assessment. Accordingly, we assumed missing at random situation.

Findings of the intention-to-treat analysis were comparable to those shown in the per protocol analysis, i.e. we did not observe any statistically significant time*group effects. Please refer to Supplementary Table 2 for results of the intention-to-treat analysis.

6.2.4.4 Differences in characteristics between positive, negative, and non-responders (intervention group, per protocol analysis)

When taking into account walking speed, stride length, and double support in all three walking conditions, between 10 % and 39 % of participants in the IG improved their gait performance by at least 10 % (considered as positive responders). Moreover, 23 % to 61 % of IG participants did not change their gait performance (considered as non-responders), while 19 % to 39 % showed a decline in gait performance by at least 10 % (considered as negative responders). Table 32 displays the proportion of positive, non-, and, negative responders in the IG depending on spatiotemporal gait parameter and walking condition, as well as mean changes in gait performance.

Table 32. Positive, non-, and negative responders in the intervention group and mean changes in gait performance (per protocol analysis)

	All		**Negative responders**		**Non-responders**		**Positive responders**	
	n	Mean change (SD)	n [%]	Mean change (SD)	n [%]	Mean change (SD)	n [%]	Mean change (SD)
Single task								
Walking speed, m/sec	89	-0.03 (0.21)	35%	-0.22 (0.09)	48%	-0.01 (0.05)	17%	0.32 (0.19)

Stride length, cm	89	-2.07 (14.98)	26%	-19.59 (10.15)	57%	-0.75 (5.42)	17%	20.30 (9.81)
Double support, % of stride time	89	3.57 (12.58)	29%	18.99 (8.62)	61%	-0.57 (4.69)	10%	-16.15 (6.08)
Dual task, counting backwards								
Walking speed, m/sec	62	0 (0.26)	39%	-0.26 (0.11)	23%	-0.02 (0.04)	39%	0.27 (0.14)
Stride length, cm	62	2.06 (17.08)	19%	-18.36 (7.05)	50%	-2.18 (4.88)	31%	21.87 (13.99)
Double support, % of stride time	62	3.16 (14.32)	29%	20.67 (9.37)	55%	-0.36 (5.56)	16%	-16.44 (4.76)
Dual task, naming animals								
Walking speed, m/sec	61	0.02 (0.29)	34%	-0.27 (0.12)	31%	-0.01 (0.06)	34%	0.33 (0.23)
Stride length, cm	61	3.49 (19.36)	23%	-18.66 (7.34)	49%	0.94 (6.27)	28%	26.21 (17.33)
Double support, % of stride time	61	-0.14 (14.68)	23%	21.12 (9.04)	56%	-1.99 (4.65)	21%	-18.20 (6.27)

n: number, SD: standard deviation

Positive, non-, and negative responders differed statistically significantly in terms of baseline performance of walking speed (both dual tasks), stride length (single task, dual task naming animals), double support (single task, both dual tasks), Timed Up & Go Test (TUG; single task: stride length, dual task naming animals: walking speed), modified 30-second chair stand test (30s CST; single task: double support), MMSE (single task), and proportion of walking aids (dual task naming animals: stride length; see Table 33, Supplementary Table 3 presents statistically non-significant results).

Table 33. Statistically significant differences in baseline motor and cognitive performance as well as the use of walking aids between positive, non-, and negative responders in the intervention group (per protocol analysis)

	Negative responders	**Non-responders**	**Positive responders**	**Between group difference**	**Post-hoc analysis**
	Mean (SD)	Mean (SD)	Mean (SD)	$F(df_{numerator}, df_{denominator})$/ $Chi^2(df)$, p	
Single task, walking speed					
MMSE (n=89)	14.8 (4.0)	18.5 (3.9)	16.7 (5.1)	$Chi^2(2)=12.093$, p=0.002	z=-3.472, p=0.002, r=0.404 [a]
Single task, stride length					
Stride length, cm (n=89)	80.3 (19.9)	89.4 (15.2)	62.7 (20.0)	F(2,86)=14.129, p<0.001, η_p^2=0.247	p=0.008, MD=-17.60, CI_{95} [-31.32, -3.89] [b] p<0.001, MD=-26.79, CI_{95} [-38.93, -14.65] [c]
TUG, sec (n=89)	22.9 (10.9)	19.0 (7.6)	31.5 (20.0)	$Chi^2(2)=8.234$, p=0.016	z=-2.800, p=0.015, r=0.325 [c]
MMSE (n=89)	14.5 (3.5)	17.8 (4.3)	17.4 (5.2)	$Chi^2(2)=9.510$, p=0.009	z=-3.046, p=0.007, r=0.354 [a]
Single task, double support					
Double support, % of stride time (n=89)	36.2 (6.9)	37.4 (7.3)	47.3 (10.4)	F(2,86)=7.721, p=0.001 η_p^2=0.152	p=0.001, MD=11.09, CI_{95} [4.13, 18.05] [b] p=0.001, MD=9.89, CI_{95} [3.41, 16.37] [c]
Modified 30s CST (n=77)	7.5 (3.3)	9.0 (3.7)	4.8 (1.7)	F(2,74)=4.508, p=0.014, η_p^2=0.109	p=0.020, MD=-4.14, CI_{95} [-7.73, -0.55] [c]
MMSE (n=89)	15.1 (3.9)	17.9 (4.3)	16.3 (5.4)	$Chi^2(2)=6.742$, p=0.034	z=-2.558, p=0.032, r=0.286 [a]
Dual task, counting backwards, walking speed					
Walking speed, m/s (n=62)	0.63 (0.17)	0.57 (0.11)	0.47 (0.15)	F(2,59)=6.336, p=0.003, η_p^2=0.177	p=0.001, MD=-15.08, CI_{95} [-25.35, -4.81] [b]
Dual task, counting backwards, stride length					
No statistically significant differences					

Dual task, counting backwards, double support						
Double support, % of stride time (n=62)	37.7 (9.2)	40.0 (7.1)	47.8 (12.9)	$Chi^2(2)=6.496$, p=0.039		z=-2.532, p=0.034, r=0.479 [b]
Dual task, naming animals, walking speed						
Walking speed, m/s (n=61)	0.52 (0.11)	0.47 (0.13)	0.36 (0.12)	F(2,58)=9.917, η_p^2=0.255	p<0.001,	p<0.001, MD=-16.40, CI_{95} [-25.44, -7.35] [b]
TUG, sec (n=61)	18.8 (9.2)	20.8 (16.9)	24.8 (12.1)	$Chi^2(2)=6.360$, p=0.042		n.s.
Dual task, naming animals, stride length						
Stride length, cm (n=61)	77.9 (14.0)	71.8 (18.3)	61.6 (18.3)	F(2,58)=3.596, η_p^2=0.110	p=0.034,	p=0.031, MD=-16.36, CI_{95} [-31.46, -1.26] [b]
Walking aid, % (n=61)	85.7 %	50.0 %	82.4 %	Chi^2=7.540, p=0.020		
Dual task, naming animals, double support						
Double support, % of stride time (n=61)	41.9 (8.6)	44.7 (7.9)	53.2 (10.2)	F(2,61)=6.570, η_p^2=0.185	p=0.003,	p=0.003, MD=11.35, CI_{95} [3.38, 19.32] [b] p=0.010, MD=8.55, CI_{95} [1.79, 15.30] [c]

30s CST: 30-second chair stand test, CI_{95}: 95 % confidence interval, df: degrees of freedom, MD: mean difference, MMSE: Mini-Mental State Examination, n: number, n.s.: not significant, SD: standard deviation, TUG: Timed Up & Go Test

[a] post-hoc analysis: statistically significant better performance of non- compared to negative responders

[b] post-hoc analysis: statistically significant worse performance of positive compared to negative responders

[c] post-hoc analysis: statistically significant worse performance of positive compared to non-responders

The post-hoc analysis (see Table 33) revealed statistically significantly 1) worse performance of positive compared to negative responders for walking speed (both dual tasks), stride length (single task, dual task naming animals), and double support (single task, both dual tasks); 2) worse performance of positive compared to non-responders for stride length (single task), double support (single task, dual task naming animals), TUG (single task: stride length), and modified 30s CST (single task: double support); as well as 3) better performance of non- compared to negative responders for MMSE (single task).

6.2.4.5 Impact of changes in underlying motor and cognitive performance on changes in gait performance (intervention group, per protocol analysis)

Several weak to moderate correlations ($|r|=0.248$-0.436, $p<0.05$) suggested relations of changes in underlying motor and cognitive performance with changes in gait performance in single and both dual task conditions. Multiple regression analyses revealed that changes in underlying motor and cognitive performance had an impact on changes in gait performance. Related models were statistically significant and explained 12.6 % to 39.4 % of the overall variance. Changes in TUG, modified 30s CST, modified Short Physical Performance Battery (SPPB), Clock Drawing Test, Digit Span forward and backward, and Trail Making Test were statistically significant regression coefficients ($p<0.05$). Table 34 presents the details of the multiple regression analysis models.

Table 34. Impacts of changes in underlying motor and cognitive performance on changes of gait performance (intervention group, per protocol analysis)

	B	SE	β	t	p	R^2	Adjusted R^2	$F(df_{numerator}, df_{denominator})$, p	f^2
Single task, changes in walking speed (n=51)									
Model						0.207	0.191	F(1,49)=12.826, p=0.001	0.24
Constant	-1.082	1.696		-0.638	0.526				
Changes in TUG	-0.915	0.256	-0.455	-3.581	0.001				
Single task, changes in stride length (n=51)									
Model						0.146	0.128	F(1,49)=8.352, p=0.006	0.15
Constant	-2.774	1.518		-1.828	0.074				
Changes in modified SPPB	2.107	0.729	0.382	2.890	0.006				
Single task, changes in double support (n=51)									
Model						0.144	0.126	F(1,49)=8.210, p=0.006	0.14
Constant	1.218	0.520		2.341	0.023				
Changes in modified SPPB	-0.716	0.250	-0.379	-2.865	0.006				
Dual task, counting backwards, changes in walking speed (n=42)									
Model						0.387	0.356	F(2,39)=12.322, p<0.001	0.55
Constant	-5.331	1.995		-2.728	0.010				
Changes in Clock Drawing Test	3.597	0.961	0.474	3.742	0.001				
Changes in modified 30s CST	2.881	0.765	0.477	3.767	0.001				
Dual task, counting backwards, changes in stride length (n=42)									
Model						0.334	0.300	F(2,39)=9.771, p<0.001	0.43
Constant	-0.661	1.626		-0.406	0.687				
Changes in modified SPPB	2.117	0.875	0.359	2.420	0.020				
Changes in modified 30s CST	1.519	0.716	0.314	2.122	0.040				

Dual task, counting backwards, changes in double support (n=42)									
Model						0.438	0.394	F(3,38)=9.871, p<0.001	0.65
Constant	2.449	0.654		3.745	0.001				
Changes in Clock Drawing Test	-1.167	0.306	-0.469	-3.814	<0.001				
Changes in modified 30s CST	-0.934	0.244	-0.472	-3.834	<0.001				
Changes in Digit Span backward	-0.682	0.333	-0.250	-2.052	0.047				
Dual task, naming animals, changes in walking speed (n=40)									
Model						0.280	0.241	F(2,37)=7.184, p=0.002	0.31
Constant	-1.991	1.823		-1.092	0.282				
Changes in modified 30s CST	1.796	0.706	0.364	2.544	0.015				
Changes in Trail Making Test	0.501	0.228	0.313	2.192	0.035				
Dual task, naming animals, changes in stride length (n=40)									
Model						0.296	0.258	F(2,37)=7.788, p=0.002	0.35
Constant	-0.360	1.681		-0.214	0.832				
Changes in modified SPPB	2.790	0.781	0.500	3.571	0.001				
Changes in Clock Drawing Test	-1.774	0.785	-0.316	-2.259	0.030				
Dual task, naming animals, changes in double support (n=40)									
Model						0.198	0.177	F(1,38)=9.378, p=0.004	0.22
Constant	1.426	0.927		1.538	0.132				
Changes in Digit Span forward	-1.736	0.567	-0.445	-3.062	0.004				

30s CST: 30-second chair stand test, df: degrees of freedom, n: number, SE: standard error, SPPB: Short Physical Performance Battery, TUG: Timed Up & Go Test

6.2.5 Discussion

6.2.5.1 Effects of the multimodal exercise program on spatiotemporal gait parameters

This multicenter randomized controlled trial aimed to investigate the effectiveness of a dementia-specific MEP, which combined motor and cognitive tasks, on gait performance. As we did not observe any statistically significant time*group effects, our primary hypothesis that a 16-week MEP, may be more effective in reducing the decline in gait performance in IWD than conventional treatment alone could not be confirmed. This may be explained by the heterogeneity of the study sample or the relatively low amount of walking tasks included in the intervention.

With regard to sample characteristics as well as motor, cognitive, and gait performance at baseline, we observed large standard deviations indicating that the sample of IWD was highly heterogeneous in this study (see Table 30, Table 31, and Supplementary Table 4). This effect may be present in many studies among IWD in general. Due to this large heterogeneity, it is very difficult to adequately tailor one standardized physical activity intervention to the needs of all participants, i.e. there is likely no standard physical activity intervention that fits all IWD.

With respect to the applied intervention, an in-depth analysis of the MEP showed that it did not include a large amount of specific walking tasks. Even though we had planned to increase the number of exercises focusing on walking throughout the intervention, this was often not possible due to our principle of ensuring the safety of participants at all times during the MEP. Additionally, we assumed that tasks aiming to improve balance, mobility, strength and function of lower limbs may be sufficient to positively affect gait performance. However, based on our findings, this assumption could not be confirmed. Thus, including a sufficient amount of specific walking exercises should be ensured in future physical activity interventions that aim at improving gait performance.

6.2.5.2 Differences in characteristics between positive, negative, and non-responders and impact of changes in underlying motor and cognitive performance on changes in gait performance

Despite not having observed positive overall effects, additional analyses showed that between 61 % and 81 % of participants in the IG improved or maintained their gait

performance after participating in the MEP. In studies among IWD, who usually experience rapid decline of motor, cognitive, and gait performance, even maintaining the current levels of performance is indicative of a beneficial effect. In order to better understand the prerequisites and impacts to induce such benefits from physical activity interventions, we conducted secondary analyses that focused on examining differences of baseline performance and sample characteristics between positive, non-, and negative responders, and also considered impacts of underlying changes in motor and cognitive performance on changes in gait performance. As compared to negative and non-responders, positive responders primarily showed lower gait performance at baseline and additionally demonstrated lower performance in single motor assessments. Moreover, non-responders were less cognitively impaired than negative responders. Accordingly, low motor and gait performance as well as increased cognitive performance seem to be prerequisites for IWD in order to benefit from the MEP. Additionally, stepwise regression analyses supported the hypothesis that changes in underlying motor and cognitive performance have an impact on changes in gait performance. Indeed, the respective statistical models explained between 12.6 % and 39.4 % of the overall variance.

Focusing on prerequisites related to the effectiveness of the MEP, the observed lower motor performance of positive responders compared to non- and negative responders at baseline may indicate a greater potential for performance improvements for participants who enter the intervention with lower baseline levels of motor performance. As described above, it was not always possible to include more complex walking tasks throughout our intervention. Accordingly, the requirements necessary to induce improvements may not have reached critical thresholds in all participants. Moreover, our findings support the assumption that IWD must have sufficient cognitive capacities in order for them to successfully participate in physical activity interventions. In contrast, severe cognitive impairments may prevent IWD from following instructions or adequately performing exercise tasks. Surprisingly, we observed a statistically significant higher cognitive performance only among non-responders and in single task conditions. Positive responders also showed higher cognitive performance than negative responders, albeit not reaching statistical significance possibly due to a relatively lower number of positive responders. When we compared cognitive performance of partici-

pants in single and dual task conditions, we observed that participants with more severe cognitive impairments were less likely to successfully perform the walking with additional dual tasks (single task: MMSE=16.9 (4.5), 45 % with MMSE<17; dual task: MMSE=18.4 (4.0)/17.8 (4.3) 29 %/26 % with MMSE<17). Accordingly, cognitive performance was more consistent in dual task conditions and did not distinguish between positive, non-, and negative responders.

Stepwise regression analyses showed different impacts of changes in underlying motor and cognitive performance, depending on spatiotemporal gait parameter and walking condition. As expected, improvements in gait performance were associated with improvements in underlying motor and cognitive performance. The observed opposite relation between stride length in dual task naming animals condition and the Clock Drawing Test requires further examination. The amount of explained variance was higher for dual task than single task conditions. In dual task conditions, changes in motor and cognitive performance were statistically significant predictors, while gait parameters in the single task condition were only affected by motor predictors. Accordingly, changes in cognitive performance may be particularly required for changes in dual task conditions, which are primarily determined by motor and cognitive demands. Dual task performance while walking is highly relevant with regard to fall prevention, and worse performance is associated with increased risk of falls (Montero-Odasso et al., 2012). Thus, fall prevention interventions should consider dual tasks and include both, motor and cognitive exercises. At the motor level, changes in strength and function of lower limbs as well as mobility were statistically significant predictors. The related performance was assessed with modified 30s CST, modified SPPB, which considers balance, mobility, and strength, and TUG. These findings indicate that there are several motor impacts related to changes in gait performance, and further emphasize the importance of multimodal interventions. Unexpectedly, changes in balance performance were not a statistically significant predictor. However, we assessed balance only in static positions, which may have different demands as compared to dynamic balance conditions while walking (Granacher, Bridenbaugh, Muehlbauer, Wehrle, & Kressig, 2011; Ringhof & Stein, 2018). Moreover, the frequent use of walking aids may have eliminated the potential impact of changes in balance performance (Schwenk et al., 2011). Assumptions at the cognitive level could not be made, as cognitive predictors differed across established regression models.

6.2.5.3 Comparison with previous studies

The findings of this randomized controlled trial are not fully in line with those observed in previous studies. In contrast to previous studies, which predominantly reported positive effects for stride length and stride time in single and dual task conditions (Coelho et al., 2013; Kemoun et al., 2010; J.-S. Kim et al., 2017; Orcioli-Silva et al., 2018; Pedrinolla et al., 2018; Perrochon et al., 2015; Schwenk, Dutzi et al., 2014; Schwenk, Zieschang et al., 2014), our investigation did not confirm the effectiveness of an MEP for these spatiotemporal gait parameters. In accordance with 20 previous studies, we did not observe statistically significant effects on walking speed (Bossers et al., 2015; Cadore et al., 2014; Dawson et al., 2019; Hageman & Thomas, 2002; Hauer et al., 2017; Junge et al., 2018; Kuiack et al., 2004; McCaffrey et al., 2014; Pedrinolla et al., 2018; Ries et al., 2015; Schwenk, Dutzi et al., 2014; Sobol et al., 2016; Souto Barreto et al., 2017; Steinberg et al., 2009; Suttanon et al., 2013; Tay et al., 2016; Telenius et al., 2015a; Thomas & Hageman, 2003; Toots et al., 2017) and percent of double support (Orcioli-Silva et al., 2018; Pedrinolla et al., 2018), while twelve others did for single (Ahn & Kim, 2015; Aman & Thomas, 2009; Bossers, Scherder et al., 2014; Hauer et al., 2012; Kemoun et al., 2010; J.-S. Kim et al., 2017; Manckoundia et al., 2014; Perrochon et al., 2015; Rolland et al., 2007; Schwenk, Zieschang et al., 2014; Toulotte et al., 2003) and dual task conditions (Tay et al., 2016). These inconsistent findings may be related to different study designs, gait assessments, interventions, and sample characteristics between previous research and our study. To the best of our knowledge, there are no published studies that compared the characteristics of positive, non-, and negative responders or investigated impacts of changes in underlying motor and cognitive performance on changes in gait performance.

6.2.5.4 Strengths and limitations

With this multicenter randomized controlled trial, we aimed at conducting high-quality research to investigate the effectiveness of a physical activity intervention on gait performance in IWD. The strengths of the study include the emphasis on high-quality methods and a detailed reporting of our methods and findings. Of note, we had a large sample size of over 300 individuals with mild to moderate dementia, our assessments were deemed adequate for IWD by an expert panel, and our MEP was specifically tailored to fit the needs and characteristics of IWD. Nevertheless, several limitations

pertain. First, multimodal interventions do not allow to unambiguously draw conclusions about causality. Observed effects may be related to the MEP itself, but could also be due to the group setting and thus enhanced social interaction, or additional attention that participants received from the exercise instructors. Additional control conditions, e.g. non-exercise groups or social visits, could have helped to limit this potential bias. Second, the assessments to determine motor performance used in this study are widely used in research but have not been specifically developed for IWD. Even though we intensively discussed the adequateness of these assessments during an expert panel (Trautwein, Barisch-Fritz et al., 2019) and carefully selected the most appropriate ones, we cannot rule out the possibility that the use of existing assessments not specifically designed for IWD may have led to biased results. For example, these assessments often do not sufficiently take into account fluctuating daily forms and motivational aspects that may play a role when examining IWD. Accordingly, results could reflect unfavorable conditions, reduced motivation, or lack of interest instead of actual motor performance. Therefore, it is critically important for future research to explore tailored motor assessments for use in IWD.

6.2.5.5 Implications

This multicenter randomized controlled trial contributes to the growing body of literature that aims at improving physical activity interventions for IWD. It shows that one standardized MEP is not effective in reducing the decline in gait performance among IWD in general. However, several participants of the IG were able to improve or maintain their gait performance after undergoing the MEP. Moreover, findings of secondary analyses allow for drawing conclusions on prerequisites and required changes that may be necessary for IWD to benefit from the MEP. These factors have important implications and should thus be considered when establishing future physical activity interventions. Our main conclusion is that it is essential to develop and provide individualized physical activity interventions for IWD, and to consider individual characteristics and needs to improve effectiveness rather than having one standardized physical activity intervention. Based on observed results in responder-non-responder-analyses we suggest tailoring physical activity interventions to baseline performance of intended outcomes and severity of cognitive impairment. To this end, we here provide preliminary criteria on how to tailor physical activity interventions to fit the specific needs of IWD. However, further investigation and refinement of these criteria is needed to better characterize

different clusters of IWD. Aside from individual characteristics and needs, it is also important to consider intended purposes when establishing physical activity interventions. Our findings indicate that physical activity interventions aiming to improve gait performance in IWD should include multimodal motor exercises (e.g. walking, strength, balance, and mobility). As changes in motor and cognitive performance are statistically significant independent predictors for changes in gait performance, it makes sense that both motor and cognitive tasks are included in interventions to potentially increase the beneficial effects on gait performance and fall prevention. Linking both conclusions, individualized approaches, which include relevant contributors for improving intended outcomes, while also tailoring requirements to prerequisites and focusing on those exercises in order to improve outcomes of especially low capacity, seem to be most promising for improving the effectiveness of physical activity interventions in IWD.

Abbreviations

30s CST: 30-second chair stand test; ANOVA: analysis of variance; CG: control group; CIRS: Cumulative Illness Rating Scale; IG: intervention group; IWD: individuals with dementia; MEP: multimodal exercise program; MMSE: Mini-Mental State Examination; SPPB: Short Physical Performance Battery; TUG: Timed Up & Go Test

Declarations

Acknowledgments

We are grateful to all participants of the study and members of our research team. We would like to thank Luisa Appelles for the coordination of fieldwork and for entering data, Bastian Anedda for establishing a script to process spatiotemporal gait parameters, Anela Hadzic for supporting the preparation of gait data, and Professor Darko Jekauc for his advice on the publishing process.
This project is financially supported by the Dietmar Hopp Foundation (St. Leon-Rot, Germany). The sponsor did not have any role in the design of the study, neither in its execution, the collection, analysis or interpretation of data, the decision to submit results nor in writing the report.

Conflict of interest/disclosure statement

The authors have no conflict of interest to report.

Supplementary material

Supplementary Table 1. Sample characteristics of participants at baseline (intention-to-treat analysis)

Supplementary Table 2. Effects of the multimodal exercise program on spatiotemporal gait parameters and dual task costs (intention-to-treat analysis)

Supplementary Table 3. Differences in baseline motor and cognitive performance as well as etiology of dementia and the use of walking aids between positive, non-, and negative responders in the intervention group (statistical nonsignificant results, per protocol analysis)

Supplementary Table 4. Heterogeneity of sample in motor and cognitive baseline performance (per protocol)

References

All references of manuscript V are included in the List of References at the end of this thesis.

7 General discussion

This thesis focuses on the effectiveness of physical activity on motor and gait performance in IWD from two perspectives: establishing a high-quality methodological approach and performing a high-quality RCT. In five research articles, we showed that

(1) there is an urgent need for tailoring motor assessments to specific characteristics of IWD and recommended appropriate motor assessments based on qualitative examination of assessments applied in previous studies (Trautwein, Barisch-Fritz et al., 2019);
(2) frequently applied motor assessments have sufficient relative reliability in IWD and recommended adequate motor assessments for identifying changes at group level based on quantitative examination of assessments applied in previous studies (Trautwein, Maurus et al., 2019);
(3) the conclusiveness of previous studies is affected by several limitations and established a high-quality study design to investigate the effectiveness of physical activity on cognitive, motor and gait performance in IWD (Trautwein et al., 2017);
(4) the effectiveness of one standardized dementia-specific MEP on motor performance is limited in IWD, probably related to large heterogeneity in this sample (Barisch-Fritz, Trautwein, Scharpf, Krell-Roesch, & Woll, submitted);
(5) a dementia-specific MEP does not have a positive effect on gait performance in all participants, as effects differ depending on initial gait performance, mobility, and lower limb strength, as well as severity of impairments in global cognition, and are affected by intervention-induced changes in lower limb strength and function, mobility, executive function, attention, and working memory (Trautwein et al., submitted).

Based on these key findings, the general discussion of this thesis focuses on the examination of the established research questions.

Research question A:

How ***high-quality studies*** need to be ***designed*** to enhance evidence concerning the ***effectiveness of physical activity*** on motor and gait performance in IWD?

(RQ_A)

Research question B:

Is ***physical activity effective*** in reducing the decline of ***motor and gait performance*** in IWD?

(RQ_B)

Aiming to answer these research questions the general discussion does only consider associated findings. Further details not directly related to research questions are intensively discussed within single manuscripts and thus are not included in the general discussion.

7.1 Establishing a high-quality methodological approach: adequateness of motor assessments applied in previous studies and example for a high-quality study design

The conclusiveness of studies examining the effectiveness of physical activity is dependent on thorough designs and sound methods. However, recent reviews consistently refer to methodological limitations of previous studies and indicate the need for further high-quality studies. Accordingly, the first aim of this thesis was to establish a high-quality methodological approach allowing to investigate the effectiveness of physical activity on motor and gait performance in IWD in a high-quality study.

Herein, the identification of adequate motor assessments was determined as a key issue (Blankevoort et al., 2010; Hauer et al., 2006; Suttanon et al., 2010). Adequate motor assessments are characterized by appropriateness for the intended population, profound psychometric properties, sensitivity to change, and standardization (Blankevoort et al., 2010; Gonçalves et al., 2018; Hauer et al., 2006; Hernández et al., 2015). Considering these requirements, manuscripts I and II performed a comprehensive examination of motor assessments in IWD by summarizing, analyzing, and discussing available findings of previous studies, aiming to answer research questions A1 and A2.

Research question A1:

Which ***motor assessments*** are ***appropriate*** for IWD based on ***qualitative*** examination?

(RQ_{A1})

Research question A2:

Which ***motor assessments*** can be ***recommended*** for IWD based on ***quantitative*** outcomes, especially psychometric properties?

(RQ_{A2})

Focusing on a qualitative examination of current research practices, manuscript I recommends a sequential approach incorporating a gradual increase from simple to more complex motor tasks and a selection of eight motor assessments including Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4 (FICSIT-4), Gro-

ningen Meander Walking Test (GMWT), TUG, 6-meter walk test (6m WT), instrumented gait analysis, modified 30-second chair stand test (30s CST), SPPB, and Physical Performance Test (PPT). Moreover, it suggests the use of external cues restricted to repeated instructions as well as the allowance of common walking devices (Trautwein, Barisch-Fritz et al., 2019). These recommendations consider cognitive impairments and generally reduced physical capacity in IWD. Additionally, they comply with the derived criteria for appropriate motor assessments including short duration, simple instructions, easy motor tasks, and adapted physical strains. Thus, recommended motor assessments seem to be the most appropriate ones of those applied in previous studies with IWD based on qualitative examination.

With respect to psychometric properties and other quantitative outcomes, manuscript II recommends Functional Reach Test (FR), GMWT, Berg Balance Scale (BBS), Performance Oriented Mobility Assessment (POMA), TUG, instrumented gait analysis assessing spatiotemporal gait parameters, STS assessments using more than one repetition, and 6min WT for IWD to assess changes of motor performance at group level. These motor assessments show sufficient relative reliability, reflect sensitivity to change determined by small to large effect sizes, and are frequently applied in previous studies. However, predominately unacceptable absolute reliability limits the identification of intra-individual changes and it needs to be considered that test-retest reliability is influenced by severity and etiology of dementia as well as external cueing (Trautwein, Maurus et al., 2019).

Comparing findings of manuscripts I and II shows that qualitative and quantitative approaches result in similar recommendations for motor assessments in IWD. GMWT, TUG, instrumented gait analysis of spatiotemporal gait parameters, and 30s CST are consistently recommended. Due to insufficient research, it was not possible to quantitatively examine FICSIT-4, SPPB, and PPT, why only qualitative appropriateness can be confirmed. No clear combined recommendations can be given for BBS, POMA, and 6min WT, as the criteria of both approaches are generally met, but some qualitative restrictions are stated. In contrast, deviations between qualitative and quantitative approaches exist for 6m WT and FR. 6m WT does not seem to be clearly adequate due to quantitative criteria and thus requires further examination, while FR is not rated appropriate for IWD considering qualitative aspects.

Further recommendations of manuscripts I and II concern the influences of severity and etiology of dementia, external cues, and walking aids on motor assessments in IWD. Two of those influences, namely, the severity of dementia and external cues, were considered in qualitative and quantitative approaches and thus are summarized in the following. There are indications that the reliability and appropriateness of some motor assessments decrease with increasing severity of dementia (Blankevoort et al., 2013; Hauer & Oster, 2008; Ries et al., 2009). Accordingly, motor assessments recommended above may especially be appropriate for individuals with mild to moderate severity of dementia, while they should be carefully applied in severe dementia. Moreover, it can be concluded that tailoring motor assessments gains in importance with increasing severity of dementia. In this context, suggesting a sequential approach is promising as it allows to adequately determine performance with different complexity levels. Besides, the qualitative approach suggested standardization of external cues allowing repeated instructions as a possible approach for tailoring motor assessments to IWD. However, quantitative examination revealed different effects of external cues dependent on motor assessments. For more complex assessments external cues may improve reliability, while the opposite association is expected for those including familiar tasks close to everyday life. Combining these findings, external cues does not necessarily enhance adequateness of motor assessments in IWD and thus are only recommended if necessary for more complex motor assessments using standardized forms such as repeated instructions.

Summarizing findings of manuscripts I and II allows recommending several motor assessments for IWD based on a comprehensive qualitative and quantitative examination. However, established recommendations are based on theoretical considerations and need to be practically evaluated. The first step for this practical evaluation was performed by applying them during our RCT investigating the effectiveness of a dementia-specific MEP. Despite supporting the general feasibility of recommended motor assessments, experiences derived from this RCT show that it is not possible to identically adopt assessment procedures commonly used in cognitively unimpaired older adults. Nonetheless, most recommended motor assessments have initially been developed for cognitively unimpaired older adults. In this context, established recommendations are not able to compensate for all challenges related to assessing motor performance in IWD but summarize the most adequate motor assessments of available ones and suggest approaches for tailoring them to IWD.

With the established recommendations, manuscripts I and II provide an important contribution to enhance the conclusiveness and standardization of motor assessments in IWD for future studies. Nevertheless, further examination is required and the need for new assessments especially developed for IWD is emphasized. Such assessments can be based on available ones, but should specifically consider characteristics of IWD. The current state of research suggests promising approaches, for example by simplifying available assessments (Bossers, van der Woude et al., 2014), differentiating between motor and cognitive requirements (Werner et al., 2018), integrating motor approaches in common cognitive tasks (Wiloth et al., 2016), considering outcomes closely related to motor performance (Schwenk, Hauer et al., 2014), or focusing on familiar tasks close to everyday life (Graessel et al., 2009).

Besides adequate motor assessments, high-quality studies require further methodological considerations, e.g. on methodological specifications and design of physical activity interventions. Related to the importance of high-quality studies for enhancing evidence, manuscript III, a study protocol for an RCT, introduces an example for a high-quality methodological approach and a dementia-specific physical activity intervention and thus contributes to answering research questions A3.

Research question A3:

How ***high-quality studies*** and ***dementia-specific physical activity interventions*** need to be ***designed*** to investigate the ***effectiveness of physical activity*** on motor and gait performance in IWD?

(RQ_{A3})

In general, high-quality studies are characterized by profound designs and methods, especially with respect to adequate motor assessments and dementia-specific physical activity interventions. Concrete criteria for high-quality studies have been derived from recent reviews. Based on these criteria and the examination of previous studies, a profound design, as well as appropriate methods, have been established in our study protocol (manuscript III; Trautwein et al., 2017). Aiming to examine accordance, Table 35 compares criteria characterizing high-quality studies in IWD (see Table 8) and characteristics determined for this RCT.

Table 35. Comparison of criteria characterizing high-quality studies and characteristics determined for this randomized controlled trial

Criteria characterizing high-quality studies (derived from Blankevoort et al., 2010; Brett et al., 2016; Farina et al., 2014; Forbes et al., 2015; Groot et al., 2016; Hauer et al., 2006; Hernández et al., 2015; Lam, Liao et al., 2018; Littbrand et al., 2011; Öhman et al., 2014; Suttanon et al., 2010)	**Characteristics determined for this randomized controlled trial** (see study protocol [manuscript III], manuscripts IV and V)
Reporting ☑	
Compliance with CONSORT statement guidelines and accurate reporting of methodological aspects, sample characteristics, and details of intervention	The study protocol as well as manuscripts IV and V follow CONSORT statement guidelines and accurately report methods (see Chapters 5.3.3, 6.1.3, 6.2.3, Figure 8, Figure 10), sample characteristics (see Table 23, Table 24, Table 30, Supplementary Table 1), and details of intervention (see Chapter 5.3.3.3).
Sample (☑)	
Sufficiently large and homogenous samples with respect to severity and etiology of dementia	With 600 participants assessed for eligibility and 319 allocated to intervention and control group, this randomized controlled trial belongs to the largest studies examining the effectiveness of physical activity in IWD (see Figure 8, Figure 10). Homogeneity of the sample could only partly be ensured by including individuals with mild to moderate severity of dementia, but no further restrictions for the etiology of dementia (see Chapter 5.3.3.2).
Outcomes and assessments ☑	
Relevant outcomes assessed with valid, reliable, and sensitive assessments appropriate for individuals with dementia, allowing a comprehensive evaluation of motor domains and cognitive subdomains	With motor and gait performance, as well as dual task ability, cognitive performance and activities of daily living, this randomized controlled trial assesses relevant outcomes. A comprehensive assessment battery including 18 motor, gait, cognitive, and activities of daily living assessments allows a comprehensive examination of related subdomains (see Chapter 5.3.3.4). The adequateness of motor assessments is discussed above with respect to research questions A1 and A2. As selected motor assessments comply with those recommended based on a comprehensive examination, it can be assumed that this randomized controlled trial meets high-quality criteria for adequate motor assessments best as possible. Due to insufficient research, some restrictions for psychometric properties exist.
Interventions ☑	
Specific interventions of sufficient duration and intensity tailored to individuals with dementia	The multimodal exercise program was specifically developed for individuals with dementia and is characterized by a didactic concept focusing on their specific needs and characteristics. It is tailored to cognitive impairments of participants and includes adjusted communication, ritualization to give orientation and familiarity, as well as adequate complexity by simple and well-structured motor and cognitive tasks (see Chapter 5.3.3.3).

	With two sessions per week over a period of 16 weeks, the duration of this randomized controlled trial complies with derivations of recent reviews suggesting physical activity two to three times per week for at least 12 to 16 weeks. Intended medium to submaximal intensity progressed throughout the intervention ensures sufficient intensity (see Chapter 5.3.3.3).
Statistical analysis ☑	
Appropriate and comprehensive presentation of data and adequate handling of missing data	Profound statistical analyses were performed and data are comprehensively presented in manuscripts IV and V. Using per protocol and intention-to-treat analyses ensures adequate handling of missing data (see Chapters 6.1.4, 6.2.4; Table 25, Table 26, Table 31, Supplementary Table 2).

CONSORT: Consolidated Standards of Reporting Trials

As shown in Table 35, characteristics of this RCT determined in our study protocol comply with derived criteria for high-quality studies. Small deviations are only observed with respect to a homogenous sample, which could not be fulfilled in order to reach a sufficient sample size. Overall, our study protocol actually gives an example of a high-quality study and thus is valuable for designing future high-quality studies.

7.2 Effectiveness of a dementia-specific multimodal exercise program

The above established methodological approach for high-quality studies provides an important basis for investigating the effectiveness of physical activity on motor and gait performance in IWD. Applying this approach, manuscripts IV and V aim to answer research questions B1 and B2.

Research question B1:

Is a dementia-specific MEP combining motor and cognitive tasks effective in ***reducing the decline of motor performance*** in IWD?

(RQ_{B1})

Research question B2:

Is a dementia-specific MEP combining motor and cognitive tasks effective in ***reducing the decline of gait performance*** in IWD?

(RQ_{B2})

Related to those research questions it has been hypothesized that a dementia-specific MEP is effective in reducing the decline of motor and gait performance in IWD. However, based on the results reported in manuscripts IV and V, associated hypotheses 1 and 2 could not be confirmed. The overall findings do not support the effectiveness of this dementia-specific MEP on motor and gait performance in IWD (Barisch-Fritz et al., submitted; Trautwein et al., submitted).

By not supporting the effectiveness of physical activity on motor and gait performance, findings of this RCT largely do not coincide with results observed in previous studies and recent reviews. Overall, they determine clear effectiveness of physical activity on balance and strength, possible effects on mobility and several spatiotemporal gait parameters during single task conditions, while the few findings available for functional performance and spatiotemporal gait parameters during dual task conditions are inconsistent. In contrast, the MEP of this RCT does not show an effect on any of these motor domains or spatiotemporal gait parameters. Due to the observed large methodologic heterogeneity, it is difficult to state concrete reasons. However, some explanatory approaches are stated below. With respect to the current state of research, it nevertheless is assumed that findings of this RCT do not support the ineffectiveness of physical activity in IWD, especially on balance, strength, mobility, and spatiotemporal gait parameters. Several influences may have compromised revealing the actual effectiveness of the MEP.

In line with this, further analyses do not allow concluding on the clear ineffectiveness of the MEP. Dependent on considered outcomes, 10 % to 40 % of the IG (per protocol sample) benefited with regard to motor and gait performance. Accordingly, there are indications for its effectiveness in some IWD, even if the overall effectiveness is limited. Relating thereto, the question for possible reasons explaining the observed limited effectiveness arises. Based on observed data structure and theoretical considerations three possible explanations were identified: i) heterogeneity of the sample, ii) fluctuating daily forms, and iii) deviations between trained and measured outcomes. Besides, critically discussing these possible influences, potential solutions are also considered in the following.

Large minimum to maximum ranges and high standard deviations in sample characteristics and outcomes (see Table 24, Table 26, Table 30, Table 31) reflect the heterogeneity of included IWD with respect to motor and cognitive performance, as well as general health conditions. Considering this heterogeneity, one consistent intervention may not be adequate for all participants. While some participants are demanded too little, others are not able to participate as intended as requirements exceed their individual capacities. Taking into account the principles of training sciences, adaptation reactions can only be achieved if a critical stimulus threshold is exceeded (Wilmore & Costill, 2004). Accordingly, it is important to consider the individual characteristics of

participants and appropriately adapt the requirements of physical activity. Thus, individualized interventions offer promising approaches to deal with the heterogeneity in IWD.

Moreover, daily forms frequently change in IWD (Deimel, Dexel, & Schreckling, 2017). Underlying components, such as motivation, general well-being, or current physical condition, may affect performance in motor assessments and conceal actual capacities. Therefore, changes in daily form between different times of measurements may be challenging. The need for motor assessments, which are robust towards such influences, again emphasizes the relevance of assessments tailored to specific characteristics of IWD. Moreover, it is important to consider daily form, when evaluating motor performance. Therefore, feasible methods to determine daily form are required.

Besides the heterogeneity of IWD and fluctuating daily forms, deviations between trained and assessed outcomes may explain the limited effectiveness of the MEP. Taking into account proportions of motor domains included in the MEP shows that strength (43 %) was the most trained motor domain. In contrast, it was insufficiently considered in the assessment battery solely applying 30s CST, which only partly depends on lower limb strength. As several other strength assessments show limited appropriateness for IWD and thus cannot be recommended based on an extensive examination, new strength assessments tailored to specific characteristics of IWD need to be developed. With the focus on the primary outcomes balance, mobility, and gait, which were considered in several assessments, the MEP includes an adequate amount of balance (25 %), while mobility and gait were not directly addressed. However, there are indications for possible effects on mobility and gait, as those motor/functional domains are closely related to balance and strength (Alexander, 1996; Bruce-Keller, Brouillette, Tudor-Locke, Foil, Gahan, Nye et al., 2012; Tiedemann et al., 2005). Nevertheless, benefits on mobility and gait performance possibly require additional incorporation of more specific components, such as tasks based on walking. As the MEP is a group physical activity intervention predominately performed in a seated position, the number of appropriate exercises may have been insufficient. Related to increased risk of falls, the selected group setting limited the application of more demanding balance tasks and exercises in a free-standing position or during walking. Reducing group size potentially allows enhancing the number of such exercises, as safety can better be ensured with increased supervision.

Despite these possible explanations and previous research supporting its effectiveness, it cannot be completely excluded that physical activity has no demonstrable effect on motor and gait performance in IWD or that it is limited to certain individuals. There may be a threshold at which further progression of the disease will prevent positive effects of physical activity. Thus, it is important to progressively extent physical activity offers to dementia prevention and also consider individuals in early and preclinical stages in future studies. Due to their lower motor and cognitive impairments, they may have a higher potential to benefit from physical activity interventions. In this way, physical activity could also have prospective impacts on dementia.

Summarizing considerations of different reasons possibly explaining the limited effectiveness of the MEP again reveals adequate assessments and dementia-specific physical activity interventions as key issues for future studies. The examination of characteristics of positive, non-, and negative responders, and impacts of changes in underlying motor and cognitive performance on changes in gait performance may result in more detailed findings, especially related to physical activity interventions.

7.3 Characteristics of positive, non-, and negative responders and impacts of changes in underlying motor and cognitive performance on changes in gait performance

Related to the limited effectiveness of the MEP, reflecting no overall effects, but several participants benefitting, additional analyses may allow concluding on characteristics of positive, non-, and negative responders and impacts of changes in underlying motor and cognitive performance on changes in gait performance. Accordingly, manuscript V performed appropriate additional analyses in order to gain deeper insight into these impacts on the effectiveness of the MEP on gait performance. With respect to research question B3, a first step is to characterize and compare positive, non-, and negative responders of the MEP.

Research question B3:

Regarding gait performance, do ***positive, non-, and negative responders*** of the MEP ***differ in specific characteristics*** closely related to gait performance?

(RQ_{B3})

Comparing motor and cognitive performance at baseline as well as etiology of dementia between positive, non-, and negative responders, reveals differences in baseline gait performance, mobility, lower limb strength, severity of cognitive impairments in

global cognition, and the use of walking aids (Trautwein et al., submitted). These findings partly confirm hypotheses H_3 and H_4. However, differences could not be shown for all considered motor and cognitive domains. They predominately exist in superordinate domains, e.g. gait performance and severity of cognitive impairments in global cognition. As presented in the theoretical foundation (see Chapter 2.2) gait is a complex process influenced by several motor and cognitive functions. Assuming individual interactions with different amounts of influences of several motor and cognitive functions, overall differences may be especially available in superordinate areas. In Contrast, hypothesis H_5, which focuses on the impact of the etiology of dementia cannot be confirmed. This may be related to the large amount of missing data and uncertainties in determined etiologies.

More detailed, positive and non-responders show lower initial gait performance and higher global cognitive performance compared to negative and/or non-responders. Based on lower gait performance observed in positive compared to negative and non-responders, it is assumed that the MEP did not provide a critical stimulus threshold for IWD with better initial gait performance. Probably, it did not contain a sufficient amount of more demanding balance and walking tasks reaching appropriate intensities for those participants. For participants with lower gait performance, these requirements, however, may have been sufficient. As mentioned above, individualization of physical activity interventions for IWD may generally enhance effectiveness as training characteristics can be specifically tailored to individual needs and prerequisites. Regarding the severity of cognitive impairments, effects on gait performance may only be possible in less advanced severities of dementia. Irreversible alterations of the brain in severe dementia, may affect potential underlying mechanisms related to impacts of physical activity. Moreover, effectively participating in group physical activity interventions requires a sufficient amount of cognitive function, e.g. allowing to follow instructions or adequately perform motor exercises. Accordingly, the MEP seems to be more adequate for mild than for moderate severities of dementia. With increasing cognitive impairments, additional and more specific adaptations may be necessary.

Besides tailoring physical activity interventions to individual needs and prerequisites, the analysis of impacts of changes in underlying motor and cognitive performance on changes in gait performance may provide further important findings for designing future

physical activity interventions specifically tailored to IWD. Accordingly, research question B4 focuses on the effects in underlying motor and cognitive performances of the MEP having an impact on gait performance.

Research question B4:

Which ***changes in underlying motor and cognitive performance*** have an ***impact on changes in gait performance*** in IWD who participated in a dementia-specific MEP?

(RQ_{B4})

Based on multiple regression analyses explaining up to 39.4 % of the variance, manuscript V shows that changes of lower limb strength and function, mobility, executive function, attention, and working memory have an impact on changes in gait performance (Trautwein et al., submitted). Thus, hypotheses H_6 and H_7 can be confirmed. It needs to be considered that observed associations differ with respect to several spatiotemporal gait parameters in single and two dual task conditions.

A detailed regard of multiple regression analyses shows that the proportion of the explained variance in models of dual task conditions is greater than in those of single task conditions. Except for changes in double support in dual task naming animals condition, all models include statistically significant motor regression coefficients. Additionally, most models for dual task conditions comprise supplementary cognitive predictors. Thus, mechanisms of changes in gait performance during dual task conditions seem to be more complex than in single task conditions and require more cognitive input. Based on the comparison of findings and hypotheses, not all theoretically assumed changes in underlying motor and cognitive performance had impacts on changes in gait performance. Motor impacts were preliminarily found for lower limb strength and function. Due to the various relations with motor performance, changes in gait performance seem to be mostly affected by changes in outcomes considering several motor domains, such as lower limb function. Moreover, changes in lower limb strength may be important predictors, as lower limb strength is of high importance for gait performance and frequently decreased in IWD. Despite its relevance for walking performance, changes in balance surprisingly did not belong to statistically significant regression coefficients. However, we only assessed static balance which may not adequately reflect dynamic balance requirements while walking (Granacher et al., 2011; Ringhof & Stein, 2018). Furthermore, numerous IWD were exposed to reduced balance requirements during walking due to the use of walking aids (Schwenk et al.,

2011). On cognitive level, we assumed that changes in executive function, attention, and working memory may have an impact on changes in gait performance. In dual task conditions, changes in all these cognitive functions are considered in related multiple regression models. However, no clear pattern of specific cognitive impacts can be observed. For more concrete conclusions, further examination is required.

Findings on the impacts of changes in underlying motor and cognitive performance on changes in gait performance are valuable for specifying exercises included in physical activity interventions. They allow to more specifically address motor and cognitive domains required to positively affect intended outcomes in IWD. Based on the findings of our high-quality RCT, physical activity interventions aiming to improve gait performance in IWD especially should focus on lower limb strength and function. Moreover, the inclusion of cognitive tasks referring to executive function, attention, and working memory is particularly advised when addressing walking in dual task conditions.

7.4 Strengths and limitations

The consideration of the aims of this thesis to establish a high-quality methodological approach and to perform an appropriate RCT as well as their scientific profound processing already emphasizes the special strengths of this work. Nevertheless, some limitations need to be stated as well. Both, strengths and limitations of the key issues of this thesis, i.e. a high-quality methodology, adequate assessments, and physical activity interventions specifically tailored to IWD, are generally discussed in the following.

With respect to establishing a high-quality methodological approach and performing an appropriate RCT, the comprehensive examination of previous studies and recent reviews, which built the basis to determine the study design and methodology, is a fundamental strength of this thesis. Moreover, the established high-quality methodological approach serves as a best-practice example and was also applied in the high-quality RCT of this thesis. Particularly noteworthy are the accurate reporting, the large sample size, adequate assessments, an intervention specifically tailored to IWD, and the comprehensive presentation of statistical analyses, which distinguish this RCT from previous studies. For organizational, time, and economic reasons, it was not possible to reach the upper level of high quality in all areas. Minor limitations refer to the retrospective trial registration, the heterogeneity of the sample with respect to etiology

of dementia and baseline performance, the inclusion of single participants without dementia diagnosis, the insufficient consideration of strength outcomes, the selection of single assessments with limited feasibility in IWD, an insufficient control for and documentation of additional influences such as other activities, a large amount of missing data, and limits of the statistical methods and used programs. The experiences gained in this RCT will contribute to overcome some of these limitations in future studies, such as a prospective trial registration, further tailoring assessments and physical activity interventions to IWD, a documentation of additional influences, an even better avoidance of missing data, or the application of appropriate programs to perform additional statistical analyses. Further limitations such as the heterogeneity of IWD, challenges in assessing the specific sample, or general conditions in nursing homes can be attributed to the specific characteristics of the target group. Given our objective of conducting research in real-life settings, these limitations may also be present in future studies.

The selection of assessments based on comprehensive considerations and discussions within the framework of an expert panel and the intensive examination of available findings from previous studies in a systematic review is a unique characteristic of this high-quality RCT and thus also belongs to the great strengths of this thesis. The emphasis on adequate assessments contributes to actually determining intended outcomes. Hereby, the chosen two-stage approach ensures that all relevant qualitative and quantitative determinants were comprehensively taken into account. Nevertheless, the established recommendations and motor assessments applied in the high-quality RCT, only comprise the most suitable ones of those used in previous RCT investigating the effectiveness of physical activity on motor performance in IWD. For some of those motor assessments, a quantitative examination even was not possible, as no appropriate investigations could be identified. Moreover, there are further motor assessments which might be highly adequate for IWD but were not considered as they were not applied in previous RCT. Accordingly, limitations concerning the established recommendations and the selection of adequate motor assessments refer to possible risk of bias caused by selective identification of potential assessments, insufficient research, and only focusing on available assessments, while not involving the development of new ones. Moreover, the recommendations for adequate motor assessments only rely on theoretical considerations that were not practically examined. In line with

this, the application of recommended motor assessments in our high-quality RCT revealed some minor limitations in feasibility. Future research is required to confirm the theoretically elaborated recommendations on adequate motor assessments.

In addition to the general high-quality methodology and the use of adequate motor assessments, the MEP specifically tailored to IWD is a feature of this high-quality RCT and one of the great strengths of this thesis. Compared to other physical activity interventions, it stands out due to its dementia-specific methodology, the combination of motor and cognitive tasks, and ritualization in the form of the imagination of experienced journeys. Furthermore, it was established based on the findings of a systematic review (Scharpf et al., 2013), which intensively analyzed previous studies aiming to determine optimal characteristics of effective physical activity interventions for IWD. Thus, the MEP can be regarded as a scientifically profound dementia-specific physical activity intervention. Another strength with respect to physical activity intervention is the comprehensive analysis of characteristics of responders and impacts of changes in underlying motor and cognitive performance on changes in gait performance. Herein, this thesis contributes to improving physical activity interventions specifically tailored to IWD. There is hardly any comparable research on determinants of the effectiveness of physical activity in IWD. Despite several strengths of the MEP and related statistical approaches, there are some limitations. The aim to specifically tailor the physical activity intervention to IWD could not be completely met. While the general needs and specific characteristics of IWD sufficiently were considered in the MEP, this was not completely possible on individual level. Within the group setting, the heterogeneity of the sample prevented individual adaptations to every single participant. Comparably, the intended progressive increase in intensity and requirements could not be implemented throughout. Thus, standardized group physical activity interventions for heterogeneous samples of IWD are of limited suitability. Appropriate countermeasures for future research refer to individualized approaches. Furthermore, limitations towards the MEP are related to the multimodality of the intervention. It cannot be clarified which of its components induces potential effects. Besides motor and cognitive tasks, social aspects of the group setting or additional attention received from the instructors possibly may have impacts. Additional control conditions, e.g. non-exercise activities or social visits, may help to control these impacts in future studies. Finally, potential risk of bias need to be considered for the additional analyses to determine

characteristics of responders and impacts of changes in underlying motor and cognitive performance on changes in gait performance due to a non-negligible amount of missing data. Taking into account the heterogeneity of the sample and various reasons for missing values, we decided only to impute missing data for primary analyses of time*group effects. As suggested above, future studies are advised to pay attention to better avoidance of missing data.

8 Conclusion and perspectives

This thesis focuses on the effectiveness of physical activity on motor and gait performance in IWD by pursuing two aims. As the conclusiveness of research is dependent on profound designs and methods, the first aim was to establish a high-quality methodological approach. Applying this approach, the second aim intended to contribute to enhancing evidence concerning the effectiveness of physical activity on motor and gait performance in IWD by performing a high-quality RCT.

The first aim, to establish a high-quality methodological approach, clearly could be achieved. Based on a comprehensive examination of current research practices as well as qualitative and quantitative determinants, recommendations for adequate motor assessments for IWD were developed. As adequate assessments are essential for examining the effectiveness of physical activity, such recommendations are highly relevant for future studies. Without appropriate, valid, reliable, sensitive, and standardized motor assessments determining evidence of the effectiveness of physical activity on motor and gait performance is not possible. Moreover, an example of a high-quality study design was elaborated, which can be useful for future studies. In establishing this high-quality methodological approach, the thesis is valuable for overcoming limitations observed in previous studies. It provides a sound basis for future research and thus contributes to improving research practices. However, there still is a need for research, especially in developing new motor assessments and physical activity interventions specifically tailored to characteristics of IWD.

The second aim, to enhance evidence by performing a high-quality RCT, could not be directly achieved. The findings of this RCT do not support the overall effectiveness of a dementia-specific MEP on motor and gait performance in IWD. As this is not in line with observations of most previous studies and recent reviews, knowledge on the effectiveness of physical activity on motor and gait performance could not be clarified. There still is inconsistency clearly indicating the need for future research. Due to limitations in previous studies and several influences potentially having affected findings in this RCT, it is not possible to draw final conclusions. Nevertheless, findings of this RCT are highly relevant for the area of research. Based on positive effects in some of the participants, it is assumed that the MEP is effective if certain prerequisites are fulfilled, despite not observing overall effects. Additional analyses provide valuable information on characteristics of responders as well as impacts of changes in underlying

motor and cognitive performance on changes in gait performance. These findings contribute to improving physical activity interventions for IWD, as they allow to tailor interventions to specific needs and characteristics of participants and specify exercises focusing on intended outcomes and underlying motor and cognitive functions.

Besides establishing a high-quality methodological approach and performing a high-quality RCT, this thesis provides important implications for research. As mentioned above, the conclusiveness of research is dependent on profound designs and sound methods. Accordingly, future studies are generally asked to attach importance to high-quality methodology. The established high-quality methodological approach, therefore, provides a concrete example. More specific implications on methodological level include the development of new adequate motor assessments and individualized physical activity interventions tailored to IWD in small groups of two to three participants. Considering established recommendations on available motor assessments, new motor assessments are especially needed for strength and endurance outcomes. Moreover, balance, mobility, and gait assessments may also benefit from further developments. Finally, the established recommendations require a practical evaluation, and feasibility as well as psychometric properties of available and newly developed motor assessments, need to be comprehensively examined. Furthermore, this thesis identified the development of physical activity interventions specifically tailored to individual characteristics and needs of IWD as a key issue for improving the effectiveness of physical activity interventions. Herein, it provides essential research perspectives on methodological level. Based on our findings on the characteristics of responders, it is assumed that the effectiveness of physical activity can be enhanced if interventions are tailored to the specific needs and characteristics of participants. In this context, individualized medicine (Hüsing, 2010) is a promising approach, which can be transferred to physical activity intervention for IWD. Individualized medicine includes the adjustment of therapeutic measures to individual disease factors. Thereby, it aims to obtain the best possible therapeutic effects of available measures (Hüsing, Hartig, Bührlen, Reiß, & Gaisser, 2008). With respect to individualized physical activity interventions, establishing interventions that are tailored to previously assessed motor and cognitive performance may similarly increase their effectiveness.

The implementation of the above-mentioned methodological research perspectives in future studies is important to enhance the evidence of the effectiveness of physical

activity on motor and gait performance in IWD. Moreover, specific research perspectives are indicated. In a first step, it is necessary to replicate the findings on characteristics of responders and impacts of changes in underlying motor and cognitive performance on changes in gait performance of this study. There is hardly any research in this field so that it is not possible to classify our results in the current state of research. Accordingly, future studies confirming and extending prerequisites and impacts support tailoring physical activity interventions to specific characteristics and needs of IWD. Moreover, further and more precise characteristics distinguishing different groups of IWD need to be identified. In this respect, adequate assessments are a decisive criterion, too. They are necessary to determine several clusters of IWD requiring different adaptations and substantive focuses concerning exercises in physical activity interventions. Future studies need to identify those motor and cognitive determinants and additional characteristics that cluster IWD best and examine, which assessments are most suitable. Another essential specific research perspective refers to the implementation and dissemination of individualized physical activity interventions. Designing effective individualized physical activity interventions requires profound knowledge in different areas, such as training sciences, course and characteristics of diseases, or methodological and didactic approaches in managing challenging behaviors. As professionals working with IWD often do not have expertise in all these areas, they need to be supported. Related thereto, technical developments of the last decades enable several perspectives in implementing and disseminating scientifically sound individualized physical activity interventions. For example, the use of mobile applications in healthcare continuously increases and shows a great variety of benefits (Mosa, Yoo, & Sheets, 2012; Ventola, 2014). In the context of implementing and disseminating individualized physical activity interventions, they, for instance, allow addressing a wide range of several target groups. Furthermore, various algorithms that facilitate the individualization of physical activity interventions can be integrated into mobile applications. They also can be promising for documentation processes, e.g. of current performance in assessments, performance development, or training documentation, which are important for the planning and control of individualized interventions. Future studies need to establish appropriate technical approaches and examine their feasibility as well as effectiveness.

Aside from methodological and specific research perspectives derived from the findings of our study, this thesis refers to some general research gaps, which need to be

examined in future studies to support, facilitate, or enable methodological and specific research approaches. First, several recent reviews indicate the need for future studies determining optimal characteristics of physical activity interventions (Brett et al., 2016; Groot et al., 2016), considering type of exercise (e.g. aerobic, strength, balance, and multimodal training alone or in combination with cognitive tasks), frequency, duration, and intensity for different outcomes (Blankevoort et al., 2010; Farina et al., 2014; Forbes et al., 2015; Karssemeijer et al., 2017). Studies comparing such characteristics of physical activity interventions are rare. The heterogeneity of previous studies hampers to derive related conclusions by comparing several studies. Various influences, for example, related to different study designs, sample characteristics, or assessment methods may have affected findings. With respect to specific research perspectives, future studies determining optimal characteristics of physical activity interventions depending on intended outcomes would be valuable for designing individualized interventions as they enable specific addressing of important determinants and underlying performances. Second, future studies should focus on underlying mechanisms explaining the effectiveness of physical activity on motor, cognitive and gait performance in IWD. Different motor and cognitive exercises may have different effects on physiologic, metabolic, and structural processes in the brain and body of IWD (Liu-Ambrose & Donaldson, 2009; Voelcker-Rehage & Niemann, 2013). In line with this, the literature discusses several possible underlying mechanisms (Burgener, Jao, Anderson, & Bossen, 2015; Gligoroska & Manchevska, 2012; Jensen, Hasselbalch, Waldemar, & Simonsen, 2015; Kennedy, Hardman, Macpherson, Scholey, & Pipingas, 2017; Lautenschlager, Cox, & Cyarto, 2012; Voelcker-Rehage & Niemann, 2013). Probably, they are related to complex, multifactorial processes (Burgener et al., 2015; Lautenschlager et al., 2012). However, research on such underlying mechanisms of physical activity in IWD is insufficient and the actual processes could not precisely be determined, yet (Burgener et al., 2015; Jensen et al., 2015; Kennedy et al., 2017; Lautenschlager et al., 2012). Considering specific research perspectives aiming to establish and improve individualized physical activity interventions for IWD, knowledge on underlying mechanisms, however, is necessary to better understand prerequisites and adaptions processes. Third, future research with respect to disease-specific knowledge is necessary. Current research perspectives in general dementia research aim to contribute to better understanding the complex syndrome of the disease (Kenigsberg et al., 2016; Khachaturian, Mesulam, Khachaturian, & Mohs, 2015; Knopman et al., 2018; Pistollato et al., 2016).

For example, they focus on research techniques, causalities of dementia, disease mechanisms, pathophysiologic processes, improvements in diagnostics, clinical syndromes, disease progression, treatment developments, and the identification of relevant outcomes (Knopman et al., 2018; Pistollato et al., 2016; Shah et al., 2016). As they provide deeper insights into the basics of the disease and allow better characterization of IWD with different severities and etiologies of dementia, they are important for specific research perspectives aiming to determine clusters of IWD with different prerequisites.

Overall, the findings of this thesis are highly relevant for the field of research, but also have essential practical implications. The established recommendations on motor assessments for IWD do not only refer to research. There are various fields for practical applications of motor assessments, such as diagnostic purposes and screening methods, or the documentation of performance developments. Most recommended motor assessments are simple to implement and do not require specific equipment. Thus, they easily can be applied in different contexts and settings. Nevertheless, it is necessary that users have basic background knowledge on assessing IWD and that importance is attached to a standardized execution and evaluation. Additionally, the established MEP specifically tailored to IWD has great practical significances as it can be used in various settings. Initially developed for institutionalized IWD, it is particularly suitable for the use in care facilities. Its application in further settings may require small adaptations. Besides providing concrete training schedules, practical implications of the MEP include the transfer of specifics characteristics of this intervention. Great potential can particular be seen in the combination of motor and cognitive tasks as well as in the ritualization through a standardized structure. Indirectly, the established high-quality methodological approach, as well as derived research perspectives, also have an important practical relevance. They are valuable for future studies and thus contribute to enhancing evidence and improving physical activity interventions for IWD. Enhanced evidence and improved interventions may facilitate access to effective physical activity offers, enabling IWD and their families to benefit directly.

The research within the framework of this thesis has triggered numerous perspectives and implications. With regard to practical relevance, the many positive feedbacks from

participants, relatives, care facilities, and training instructors show that these are successfully implemented[14]:

> “The exercise program is varied and amusing. It's fun to participate.”
>
> Participant of the MEP

> “It is a pleasure to see how the participants improve and can remember more and more.”
>
> Caregiver

> “We are currently offering weekly exercise interventions of imagined journeys [...]. The offer is very well received. Most of the at least ten residents take part every week. [...] In any case, the exercise interventions of imagined journeys have become a great enrichment of our activity program and many ideas that we know from you serve as inspiration.”
>
> Employee of a care facility

Such positive feedback is worth every effort associated with high-quality studies and the greatest motivation to continue exploring the effectiveness of physical activity in IWD in the future.

[14] Original quotation in German, freely translated.

List of references

Abernethy, B. (1988). Dual-task methodology and motor skills research: Some applications and methodological constraints. *Journal of Human Movement Studies*, *14*(3), 101–132.

Aguiar, P., Monteiro, L., Feres, A., Gomes, I., & Melo, A. (2014). Rivastigmine transdermal patch and physical exercises for Alzheimer's disease: A randomized clinical trial. *Current Alzheimer Research*, *11*(6), 532–537. https://doi.org/10.2174/1567205011666140618102224

Ahlskog, J. E., Geda, Y. E., Graff-Radford, N. R., & Petersen, R. C. (2011). Physical exercise as a preventive or disease-modifying treatment of dementia and brain aging. *Mayo Clinic Proceedings*, *86*(9), 876–884. https://doi.org/10.4065/mcp.2011.0252

Ahn, N., & Kim, K. [Kijin] (2015). Effects of an elastic band resistance exercise program on lower extremity muscle strength and gait ability in patients with Alzheimer's disease. *Journal of Physical Therapy Science*, *27*(6), 1953–1955. https://doi.org/10.1589/jpts.27.1953

Albert, M. S. (2011). Changes in cognition. *Neurobiology of Aging*, *32*(Suppl 1), S58-63. https://doi.org/10.1016/j.neurobiolaging.2011.09.010

Alencar, M. A., Dias, J. M. D., Figueiredo, L. C., & Dias, R. C. (2012). Handgrip strength in elderly with dementia: Study of reliability. *Brazilian Journal of Physical Therapy*, *16*(6), 510–514. https://doi.org/10.1590/S1413-35552012005000059

Alexander, N. B. (1996). Gait disorders in older adults. *Journal of the American Geriatrics Society*, *44*(4), 434–451. https://doi.org/10.1111/j.1532-5415.1996.tb06417.x

Alexander, N. B., & Goldberg, A. (2005). Gait disorders: Search for multiple causes. *Cleveland Clinic Journal of Medicine*, *72*(7), 586-600. https://doi.org/10.3949/ccjm.72.7.586

Alexander, N. B., & Hausdorff, J. M. (2008). Guest editorial: Linking thinking, walking, and falling. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *63*(12), 1325–1328. https://doi.org/10.1093/gerona/63.12.1325

Allali, G., Annweiler, C., Blumen, H. M., Callisaya, M. L., Cock, A.-M. de, Kressig, R. W., . . . Beauchet, O. (2016). Gait phenotype from mild cognitive impairment to moderate dementia: Results from the GOOD initiative. *European Journal of Neurology*, *23*(3), 527–541. https://doi.org/10.1111/ene.12882

Allali, G., & Verghese, J. (2017). Management of gait changes and fall risk in MCI and dementia. *Current Treatment Options in Neurology*, *19*(29). https://doi.org/10.1007/s11940-017-0466-1

Allan, L. M., Ballard, C. G., Burn, D. J., & Kenny, R. A. (2005). Prevalence and severity of gait disorders in Alzheimer's and non-Alzheimer's dementias. *Journal of the American Geriatrics Society*, *53*(10), 1681–1687. https://doi.org/10.1111/j.1532-5415.2005.53552.x

Allan, L. M., Ballard, C. G., Rowan, E. N., & Kenny, R. A. (2009). Incidence and prediction of falls in dementia: A prospective study in older people. *PloS One*, *4*(5), e5521. https://doi.org/10.1371/journal.pone.0005521

Alzheimer’s Disease International (2015). World Alzheimer Report 2015 - The global impact of dementia: An analysis of prevalence, incidence, cost and trends. Retrieved from https://www.alz.co.uk/research/WorldAlzheimerReport2015.pdf

Alzheimer’s Disease International (2018). World Alzheimer Report 2018 - The state of the art of dementia research: New frontiers. Retrieved from https://www.alz.co.uk/research/WorldAlzheimerReport2018.pdf

Alzheimer’s Disease International (2019). World Alzheimer Report 2019: Attitudes to dementia. Retrieved from https://www.alz.co.uk/research/WorldAlzheimerReport2019.pdf

Alzheimer's Association (2019). 2019 Alzheimer's disease facts and figures. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *15*(3), 321–387. https://doi.org/10.1016/j.jalz.2019.01.010

Aman, E., & Thomas, D. R. (2009). Supervised exercise to reduce agitation in severely cognitively impaired persons. *Journal of the American Medical Directors Association*, *10*(4), 271–276. https://doi.org/10.1016/j.jamda.2008.12.053

Amboni, M., Barone, P., & Hausdorff, J. M. (2013). Cognitive contributions to gait and falls: Evidence and implications. *Movement Disorders: Official Journal of the Movement Disorder Society*, *28*(11), 1520–1533. https://doi.org/10.1002/mds.25674

American Geriatrics Society, British Geriatrics Society, & American Academy of Orthopaedic Surgeons Panel on Falls Prevention (2001). Guideline for the prevention of falls in older persons. *Journal of the American Geriatrics Society*, *49*(5), 664–672. https://doi.org/10.1046/j.1532-5415.2001.49115.x

Angevaren, M., Aufdemkampe, G., Verhaar, H. J. J., Aleman, A., & Vanhees, L. (2008). Physical activity and enhanced fitness to improve cognitive function in older people without known cognitive impairment. *The Cochrane Database of Systematic Reviews.* Advance online publication. https://doi.org/10.1002/14651858.CD005381.pub3

Annweiler, C., Beauchet, O., Celle, S., Roche, F., Annweiler, T., Allali, G., . . . Montero-Odasso, M. (2012). Contribution of brain imaging to the understanding of gait disorders in Alzheimer's disease: A systematic review. *American Journal of Alzheimer's Disease and Other Dementias*, *27*(6), 371–380. https://doi.org/10.1177/1533317512454710

Arcoverde, C., Deslandes, A., Moraes, H., Almeida, C., Araujo, N. B. d., Vasques, P. E., . . . Laks, J. (2014). Treadmill training as an augmentation treatment for Alzheimer's disease: A pilot randomized controlled study. *Arquivos De Neuro-Psiquiatria*, *72*(3), 190–196. https://doi.org/10.1590/0004-282X20130231

Arshinta, I. N., Fitriana, L. A., Adikusuma, T., Rohaedi, S., & Putri, T. P. (2018). Cardiorespiratory endurance and balance of elderly dementia patients in nursing home. *IOP Conference Series. Materials Science and Engineering*, *288.* https://doi.org/10.1088/1757-899X/288/1/012084

Ashton-Miller, J. A. (2005). Age-associated changes in the biomechanics of gait and gait-related falls in older adults. In J. M. Hausdorff & N. B. Alexander (Eds.), *Gait disorders: Evaluation and management* (pp. 63–100). Boca Raton: Taylor & Francis. https://doi.org/10.1201/b14109-5

Astrand, P. O., & Ryhming, I. (1954). A nomogram for calculation of aerobic capacity (physical fitness) from pulse rate during sub-maximal work. *Journal of Applied Physiology*, *7*(2), 218–221. https://doi.org/10.1152/jappl.1954.7.2.218

Attems, J., & Jellinger, K. A. (2014). The overlap between vascular disease and Alzheimer's disease - lessons from pathology. *BMC Medicine*, *12*, 206. https://doi.org/10.1186/s12916-014-0206-2

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews. Neuroscience*, *4*(10), 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A., Logie, R., Bressi, S., Della Sala, S., & Spinnler, H. (1986). Dementia and working memory. *The Quarterly Journal of Experimental Psychology Section a*, *38*(4), 603–618. https://doi.org/10.1080/14640748608401616

Barisch-Fritz, B., Trautwein, S., Scharpf, A., Krell-Roesch, J., & Woll, A. (submitted). Effects of a 16-week multimodal exercise program on motor performance in individuals with dementia: A multicenter randomized controlled trial.

Bean, J. F., Kiely, D. K., LaRose, S., Alian, J., & Frontera, W. R. (2007). Is stair climb power a clinically relevant measure of leg power impairments in at-risk older adults? *Archives of Physical Medicine and Rehabilitation*, *88*(5), 604–609. https://doi.org/10.1016/j.apmr.2007.02.004

Beauchet, O., Allali, G., Berrut, G., Hommet, C., Dubost, V., & Assal, F. (2008). Gait analysis in demented subjects: Interests and perspectives. *Neuropsychiatric Disease and Treatment*, *4*(1), 155–160. https://doi.org/10.2147/ndt.s2070

Beauchet, O., Allali, G., Launay, C. P., Herrmann, F. R., & Annweiler, C. (2013). Gait variability at fast-pace walking speed: A biomarker of mild cognitive impairment? *The Journal of Nutrition, Health & Aging*, *17*(3), 235–239. https://doi.org/10.1007/s12603-012-0394-4

Beauchet, O., Blumen, H. M., Callisaya, M. L., Cock, A.-M. de, Kressig, R. W., Srikanth, V. K., . . . Allali, G. (2018). Spatiotemporal gait characteristics associated with cognitive impairment: A multicenter cross-sectional study, the intercontinental "Gait, cOgnitiOn & Decline" Initiative. *Current Alzheimer Research*, *15*(3), 273–282. https://doi.org/10.2174/1567205014666170725125621

Beauchet, O., Launay, C. P., Sekhon, H., Montembeault, M., & Allali, G. (2019). Association of hippocampal volume with gait variability in pre-dementia and dementia stages of Alzheimer disease: Results from a cross-sectional study. *Experimental Gerontology*, *115*, 55–61. https://doi.org/10.1016/j.exger.2018.11.010

Bennett, D. A., Schneider, J. A., Buchman, A. S., Barnes, L. L., Boyle, P. A., & Wilson, R. S. (2012). Overview and findings from the rush Memory and Aging Project. *Current Alzheimer Research*, *9*(6), 646–663. https://doi.org/10.2174/156720512801322663

Berg, K. (1989). Measuring balance in the elderly: Preliminary development of an instrument. *Physiotherapy Canada*, *41*(6), 304–311. https://doi.org/10.3138/ptc.41.6.304

Berr, C., Wancata, J., & Ritchie, K. (2005). Prevalence of dementia in the elderly in Europe. *European Neuropsychopharmacology: The Journal of the European College of Neuropsychopharmacology*, *15*(4), 463–471. https://doi.org/10.1016/j.euroneuro.2005.04.003

Bessou, P., Dupui, P., Montoya, R., & Pagès, B. (1988). Simultaneous recording of longitudinal displacements of both feet during human walking. *Journal De Physiologie*, *83*(2), 102–110.

Blankevoort, C. G., van Heuvelen, M. J. G., Boersma, F., Luning, H., Jong, J. de, & Scherder, E. J. A. (2010). Review of effects of physical activity on strength, balance, mobility and ADL performance in elderly subjects with dementia. *Dementia and Geriatric Cognitive Disorders*, *30*(5), 392–402. https://doi.org/10.1159/000321357

Blankevoort, C. G., van Heuvelen, M. J. G., & Scherder, E. J. A. (2013). Reliability of six physical performance tests in older people with dementia. *Physical Therapy*, *93*(1), 69–78. https://doi.org/10.2522/ptj.20110164

Bohannon, R. W. (1995). Sit-to-stand test for measuring performance of lower extremity muscles. *Perceptual and Motor Skills*, *80*(1), 163–166. https://doi.org/10.2466/pms.1995.80.1.163

Bossers, W. J. R., Scherder, E. J. A., Boersma, F., Hortobágyi, T., van der Woude, L. H. V., & van Heuvelen, M. J. G. (2014). Feasibility of a combined aerobic and strength training program and its effects on cognitive and physical function in institutionalized dementia patients. A pilot study. *PloS One*, *9*(5), e97577. https://doi.org/10.1371/journal.pone.0097577

Bossers, W. J. R., van der Woude, L. H. V., Boersma, F., Hortobágyi, T., Scherder, E. J. A., & van Heuvelen, M. J. G. (2015). A 9-week aerobic and strength training program improves cognitive and motor function in patients with dementia: A randomized, controlled trial. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, *23*(11), 1106–1116. https://doi.org/10.1016/j.jagp.2014.12.191

Bossers, W. J. R., van der Woude, L. H. V., Boersma, F., Hortobágyi, T., Scherder, E. J. A., & van Heuvelen, M. J. G. (2016). Comparison of effect of two exercise programs on activities of daily living in individuals with dementia: A 9-week randomized, controlled trial. *Journal of the American Geriatrics Society*, *64*(6), 1258–1266. https://doi.org/10.1111/jgs.14160

Bossers, W. J. R., van der Woude, L. H. V., Boersma, F., Scherder, E. J. A., & van Heuvelen, M. J. G. (2012). Recommended measures for the assessment of cognitive and physical performance in older patients with dementia: A systematic review. *Dementia and Geriatric Cognitive Disorders Extra*, *2*(1), 589–609. https://doi.org/10.1159/000345038

Bossers, W. J. R., van der Woude, L. H. V., Boersma, F., Scherder, E. J. A., & van Heuvelen, M. J. G. (2014). The Groningen Meander Walking Test: A dynamic walking test for older adults with dementia. *Physical Therapy*, *94*(2), 262–272. https://doi.org/10.2522/ptj.20130077

Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., & Ravaud, P. (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: Explanation and elaboration. *Annals of Internal Medicine*, *148*(4), 295–309. https://doi.org/10.7326/0003-4819-148-4-200802190-00008

Brach, J. S., FitzGerald, S., Newman, A. B., Kelsey, S., Kuller, L. H., VanSwearingen, J. M., & Kriska, A. M. (2003). Physical activity and functional status in community-dwelling older women: A 14-year prospective study. *Archives of Internal Medicine*, *163*(21), 2565–2571. https://doi.org/10.1001/archinte.163.21.2565.

Brett, L., Traynor, V., & Stapley, P. (2016). Effects of physical exercise on health and well-being of individuals living with a dementia in nursing homes: A systematic review. *Journal of the American Medical Directors Association*, *17*(2), 104–116. https://doi.org/10.1016/j.jamda.2015.08.016

Bridenbaugh, S. A., & Kressig, R. W. (2015). Motor cognitive dual tasking: Early detection of gait impairment, fall risk and cognitive decline. *Zeitschrift Für Gerontologie Und Geriatrie*, *48*(1), 15–21. https://doi.org/10.1007/s00391-014-0845-0

Bronas, U. G., Salisbury, D., Kelly, K., Leon, A., Chow, L., & Yu, F. (2017). Determination of aerobic capacity via cycle ergometer exercise testing in Alzheimer's disease. *American Journal of Alzheimer's Disease and Other Dementias*, *32*(8), 500–508. https://doi.org/10.1177/1533317517720065

Bruce-Keller, A. J., Brouillette, R. M., Tudor-Locke, C., Foil, H. C., Gahan, W. P., Correa, J., . . . Keller, J. N. (2012). Assessment of cognition, physical performance, and gait in the context of mild cognitive impairment and dementia. *Journal of the American Geriatrics Society*, *60*(1), 176–177. https://doi.org/10.1111/j.1532-5415.2011.03762.x

Bruce-Keller, A. J., Brouillette, R. M., Tudor-Locke, C., Foil, H. C., Gahan, W. P., Nye, D. M., . . . Keller, J. N. (2012). Relationship between cognitive domains, physical performance, and gait in elderly and demented subjects. *Journal of Alzheimer's Disease: JAD*, *30*(4), 899–908. https://doi.org/10.3233/JAD-2012-120025

Brunnström, H., Gustafson, L., Passant, U., & Englund, E. (2009). Prevalence of dementia subtypes: A 30-year retrospective survey of neuropathological reports. *Archives of Gerontology and Geriatrics*, *49*(1), 146–149. https://doi.org/10.1016/j.archger.2008.06.005

Bruton, A., Conway, J. H., & Holgate, S. T. (2000). Reliability: What is it, and how is it measured? *Physiotherapy*, *86*(2), 94–99. https://doi.org/10.1016/S0031-9406(05)61211-4

Buchman, A. S., & Bennett, D. A. (2011). Loss of motor function in preclinical Alzheimer's disease. *Expert Review of Neurotherapeutics*, *11*(5), 665–676. https://doi.org/10.1586/ern.11.57

Buracchio, T., Dodge, H. H., Howieson, D., Wasserman, D., & Kaye, J. (2010). The trajectory of gait speed preceding mild cognitive impairment. *Archives of Neurology*, *67*(8), 980–986. https://doi.org/10.1001/archneurol.2010.159

Burgener, S. C., Jao, Y.-L., Anderson, J. G., & Bossen, A. L. (2015). Mechanism of action for nonpharmacological therapies for individuals with dementia: Implications for practice and research. *Research in Gerontological Nursing*, *8*(5), 240–259. https://doi.org/10.3928/19404921-20150429-02

Burgener, S. C., Yang, Y., Gilbert, R., & Marsh-Yant, S. (2008). The effects of a multimodal intervention on outcomes of persons with early-stage dementia. *American Journal of Alzheimer's Disease and Other Dementias*, *23*(4), 382–394. https://doi.org/10.1177/1533317508317527

Burns, J. M., Galvin, J. E., Roe, C. M., Morris, J. C., & McKeel, D. W. (2005). The pathology of the substantia nigra in Alzheimer disease with extrapyramidal signs. *Neurology*, *64*(8), 1397–1403. https://doi.org/10.1212/01.WNL.0000158423.05224.7F

Cadore, E. L., Casas-Herrero, A., Zambom-Ferraresi, F., Martínez-Ramírez, A., Millor, N., Gómez, M., . . . Izquierdo, M. (2015). Do frailty and cognitive impairment affect dual-task cost during walking in the oldest old institutionalized patients? *Age (Dordrecht, Netherlands)*, *37*(6), 124. https://doi.org/10.1007/s11357-015-9862-1

Cadore, E. L., Moneo, A. B. B., Mensat, M. M., Muñoz, A. R., Casas-Herrero, A., Rodriguez-Mañas, L., & Izquierdo, M. (2014). Positive effects of resistance training in frail elderly patients with dementia after long-term physical restraint. *Age (Dordrecht, Netherlands)*, *36*(2), 801–811. https://doi.org/10.1007/s11357-013-9599-7

Callisaya, M. L., Blizzard, L., McGinley, J. L., Schmidt, M. D., & Srikanth, V. K. (2010). Sensorimotor factors affecting gait variability in older people - a population-based study. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *65*(4), 386–392. https://doi.org/10.1093/gerona/glp184

Callisaya, M. L., Blizzard, L., Schmidt, M. D., McGinley, J. L., Lord, S. R., & Srikanth, V. K. (2009). A population-based study of sensorimotor factors affecting gait in older people. *Age and Ageing*, *38*(3), 290–295. https://doi.org/10.1093/ageing/afp017

Camicioli, R., Howieson, D., Oken, B., Sexton, G., & Kaye, J. (1998). Motor slowing precedes cognitive impairment in the oldest old. *Neurology*, *50*(5), 1496–1498. https://doi.org/10.1212/WNL.50.5.1496

Cancela, J. M., Ayán, C., Varela, S., & Seijo, M. (2016). Effects of a long-term aerobic exercise intervention on institutionalized patients with dementia. *Journal of Science and Medicine in Sport*, *19*(4), 293–298. https://doi.org/10.1016/j.jsams.2015.05.007

Carter, R., Lubinsky, J., & Domholdt, E. (2013). *Rehabilitation research: Principles and applications* (4. ed.): Elsevier Health Sciences.

Caspersen, C. J., Powell, K. E., & Christenson, G. M. (1985). Physical activity, exercise, and physical fitness: Definitions and distinctions for health-related research. *Public Health Reports (Washington, D.C.: 1974)*, *100*(2), 126–131.

Castellani, R. J., Lee, H.-G. [Hyoung-Gon], Zhu, X., Perry, G., & Smith, M. A. (2008). Alzheimer disease pathology as a host response. *Journal of Neuropathology and Experimental Neurology*, *67*(6), 523–531. https://doi.org/10.1097/NEN.0b013e318177eaf4

Chan, A.-W., Tetzlaff, J. M., Altman, D. G., Laupacis, A., Gøtzsche, P. C., Krleza-Jeric, K., . . . Moher, D. (2013). Spirit 2013 statement: Defining standard protocol items for clinical trials. *Annals of Internal Medicine*, *158*(3), 200–207. https://doi.org/10.7326/0003-4819-158-3-201302050-00583

Chan, A.-W., Tetzlaff, J. M., Gøtzsche, P. C., Altman, D. G., Mann, H., Berlin, J. A., . . . Moher, D. (2013). Spirit 2013 explanation and elaboration: Guidance for protocols of clinical trials. *BMJ (Clinical Research Ed.)*, *346*, e7586. https://doi.org/10.1136/bmj.e7586

Chew-Graham, C., & Ray, M. (Eds.) (2016). *Mental health and older people: A guide for primary care practitioners*. Switzerland: Springer. https://doi.org/10.1007/978-3-319-29492-6_16

Chodzko-Zajko, W. J., Proctor, D. N., Fiatarone Singh, M. A., Minson, C. T., Nigg, C. R., Salem, G. J., & Skinner, J. S. (2009). American College of Sports Medicine position stand. Exercise and physical activity for older adults. *Medicine and*

Science in Sports and Exercise, *41*(7), 1510–1530. https://doi.org/10.1249/MSS.0b013e3181a0c95c

Choi, J.-S., Oh, H.-S., Kang, D.-W., Mun, K.-R., Choi, M.-H., Lee, S.-J. [Su-Jeong], . . . Tack, G.-R. (2011). Comparison of gait and cognitive function among the elderly with Alzheimer's disease, mild cognitive impairment and healthy. *International Journal of Precision Engineering and Manufacturing*, *12*(1), 169–173. https://doi.org/10.1007/s12541-011-0024-9

Christofoletti, G., Oliani, M. M., Gobbi, S., Stella, F., Bucken Gobbi, L. T., & Renato Canineu, P. (2008). A controlled clinical trial on the effects of motor intervention on balance and cognition in institutionalized elderly patients with dementia. *Clinical Rehabilitation*, *22*(7), 618–626. https://doi.org/10.1177/0269215507086239

Clark, R. A., Bryant, A. L., Pua, Y., McCrory, P., Bennell, K., & Hunt, M. (2010). Validity and reliability of the Nintendo Wii Balance Board for assessment of standing balance. *Gait & Posture*, *31*(3), 307–310. https://doi.org/10.1016/j.gaitpost.2009.11.012

Coelho, F. G. d. M., Andrade, L. P., Pedroso, R. V., Santos-Galduróz, R. F., Gobbi, S., Costa, J. L. R., & Gobbi, L. T. B. (2013). Multimodal exercise intervention improves frontal cognitive functions and gait in Alzheimer's disease: A controlled trial. *Geriatrics & Gerontology International*, *13*(1), 198–203. https://doi.org/10.1111/j.1447-0594.2012.00887.x

Coelho, F. G. d. M., Stella, F., Andrade, L. P., Barbieri, F. A., Santos-Galduróz, R. F., Gobbi, S., . . . Gobbi, L. T. B. (2012). Gait and risk of falls associated with frontal cognitive functions at different stages of Alzheimer's disease. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *19*(5), 644–656. https://doi.org/10.1080/13825585.2012.661398

Cohen, J. [Jacob] (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J. [Jacob] (1988). *Statistical power analysis for the behavioral sciences* (2 edition). Hillsdale, N.J.: Lawrence Erlbaum.

Cohen, J. A., Verghese, J., & Zwerling, J. L. (2016). Cognition and gait in older people. *Maturitas*, *93*, 73–77. https://doi.org/10.1016/j.maturitas.2016.05.005

Cohen-Mansfield, J. (2000). Heterogeneity in dementia: Challenges and opportunities. *Alzheimer Disease and Associated Disorders*, *14*(2), 60–63. https://doi.org/10.1097/00002093-200004000-00002

Cooper, K. H. (1968). A means of assessing maximal oxygen intake. *JAMA*, *203*(3), 201. https://doi.org/10.1001/jama.1968.03140030033008

Cott, C. A., Dawson, P., Sidani, S., & Wells, D. (2002). The effects of a walking/talking program on communication, ambulation, and functional status in residents with Alzheimer disease. *Alzheimer Disease and Associated Disorders*, *16*(2), 81–87. https://doi.org/10.1097/00002093-200204000-00005

Csuka, M., & McCarty, D. J. (1985). Simple method for measurement of lower extremity muscle strength. *The American Journal of Medicine*, *78*(1), 77–81. https://doi.org/10.1016/0002-9343(85)90465-6

Dawson, N., Judge, K. S., & Gerhart, H. (2019). Improved functional performance in individuals with dementia after a moderate-intensity home-based exercise program:

A randomized controlled trial. *Journal of Geriatric Physical Therapy (2001)*, *42*(1), 18–27. https://doi.org/10.1519/JPT.0000000000000128

Deimel, H., Dexel, M., & Schreckling, S. (2017). Bewegung und körperliche Aktivität im ambulanten Setting bei Demenzerkrankungen. *Erfahrungsheilkunde*, *66*(2), 108–114. https://doi.org/10.1055/s-0043-106402

Desmond, D. W. (2004a). The neuropsychology of vascular cognitive impairment: Is there a specific cognitive deficit? *Journal of the Neurological Sciences*, *226*(1-2), 3–7. https://doi.org/10.1016/j.jns.2004.09.002

Desmond, D. W. (2004b). Vascular dementia. *Clinical Neuroscience Research*, *3*(6), 437–448. https://doi.org/10.1016/j.cnr.2004.04.009

Di Santo, S. G., Prinelli, F., Adorni, F., Caltagirone, C., & Musicco, M. (2013). A meta-analysis of the efficacy of donepezil, rivastigmine, galantamine, and memantine in relation to severity of Alzheimer's disease. *Journal of Alzheimer's Disease: JAD*, *35*(2), 349–361. https://doi.org/10.3233/JAD-122140

Dorfman, M., Mirelman, A., Hausdorff, J. M., & Giladi, N. (2014). Gait disorders in patients with cognitive impairment or dementia. In M. Merello & S. E. Starkstein (Eds.), *Movement disorders in dementias* (pp. 17–44). London: Springer. https://doi.org/10.1007/978-1-4471-6365-7_2

Du, Z., Li, Y. [Yuewei], Li, J. [Jinwei], Zhou, C., Li, F., & Yang, X. (2018). Physical activity can improve cognition in patients with Alzheimer's disease: A systematic review and meta-analysis of randomized controlled trials. *Clinical Interventions in Aging*, *13*, 1593–1603. https://doi.org/10.2147/CIA.S169565

Duncan, P. W., Weiner, D. K., Chandler, J., & Studenski, S. A. (1990). Functional reach: A new clinical measure of balance. *Journal of Gerontology*, *45*(6), M192-M197. https://doi.org/10.1093/geronj/45.6.M192

Dyer, S. M., Harrison, S. L., Laver, K., Whitehead, C., & Crotty, M. (2018). An overview of systematic reviews of pharmacological and non-pharmacological interventions for the treatment of behavioral and psychological symptoms of dementia. *International Psychogeriatrics*, *30*(3), 295–309. https://doi.org/10.1017/S1041610217002344

Eggermont, L. H., Gavett, B. E., Volkers, K. M., Blankevoort, C. G., Scherder, E. J. A., Jefferson, A. L., . . . Stern, R. A. (2010). Lower-extremity function in cognitively healthy aging, mild cognitive impairment, and Alzheimer's disease. *Archives of Physical Medicine and Rehabilitation*, *91*(4), 584–588. https://doi.org/10.1016/j.apmr.2009.11.020

Elwood, R. W. (1995). The California Verbal Learning Test: Psychometric characteristics and clinical application. *Neuropsychology Review*, *5*(3), 173–201. https://doi.org/10.1007/bf02214761

Enright, P. L. (2003). The six-minute walk test. *Respiratory Care*, *48*(8), 783–785.

Fabre, C., Chamari, K., Mucci, P., Masse-Biron, J., & Prefaut, C. (2002). Improvement of cognitive function by mental and/or individualized aerobic training in healthy elderly subjects. *International Journal of Sports Medicine*, *23*(6), 415–421. https://doi.org/10.1055/s-2002-33735

Fajersztajn, L., Cordeiro, R. C., Andreoni, S., & Garcia, J. T. (2008). Effects of functional physical activity on the maintenance of motor function in Alzheimer's disease.

Dementia & Neuropsychologia, *2*(3), 233–240. https://doi.org/10.1590/S1980-57642009DN20300013

Farina, N., Rusted, J., & Tabet, N. (2014). The effect of exercise interventions on cognitive outcome in Alzheimer's disease: A systematic review. *International Psychogeriatrics*, *26*(1), 9–18. https://doi.org/10.1017/S1041610213001385

Farrell, M. K., Rutt, R. A., Lusardi, M. M., & Williams, A. K. (2010). Reliability of the Physical Performance Test in people with dementia. *Physical & Occupational Therapy in Geriatrics*, *28*(2), 144–153. https://doi.org/10.3109/02703181.2010.487973

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feldman, H. H., & Woodward, M. (2005). The staging and assessment of moderate to severe Alzheimer disease. *Neurology*, *65*(6 Suppl 3), S10-S17. https://doi.org/10.1212/WNL.65.6_suppl_3.S10

Ferrer, I. (2010). Cognitive impairment of vascular origin: Neuropathology of cognitive impairment of vascular origin. *Journal of the Neurological Sciences*, *299*(1-2), 139–149. https://doi.org/10.1016/j.jns.2010.08.039

Fogassi, L., & Luppino, G. (2005). Motor functions of the parietal lobe. *Current Opinion in Neurobiology*, *15*(6), 626–631. https://doi.org/10.1016/j.conb.2005.10.015

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

Forbes, D., Forbes, S. C., Blake, C. M., Thiessen, E. J., & Forbes, S. [Sean] (2015). Exercise programs for people with dementia. *The Cochrane Database of Systematic Reviews*. (4), CD006489. https://doi.org/10.1002/14651858.CD006489.pub4

Fox, B., Henwood, T., Keogh, J., & Neville, C. (2016). Psychometric viability of measures of functional performance commonly used for people with dementia: A systematic review of measurement properties. *JBI Database of Systematic Reviews and Implementation Reports*, *14*(8), 115–171. https://doi.org/10.11124/JBISRIR-2016-003064

Fox, B., Henwood, T., Neville, C., & Keogh, J. (2014). Relative and absolute reliability of functional performance measures for adults with dementia living in residential aged care. *International Psychogeriatrics*, *26*(10), 1659–1667. https://doi.org/10.1017/S1041610214001124

Francese, T., Sorrell, J., & Butler, F. R. (1997). The effects of regular exercise on muscle strength and functional abilities of late stage Alzheimer's residents. *American Journal of Alzheimer's Disease and Other Dementias*, *12*(3), 122–127. https://doi.org/10.1177/153331759701200305

Freiberger, E., Vreede, P. de, Schoene, D., Rydwik, E., Mueller, V., Frändin, K., & Hopman-Rock, M. (2012). Performance-based physical function in older community-dwelling persons: A systematic review of instruments. *Age and Ageing*, *41*(6), 712–721. https://doi.org/10.1093/ageing/afs099

Geldmacher, D. S. (2003). Visuospatial dysfunction in the neurodegenerative diseases. *Frontiers in Bioscience: A Journal and Virtual Library*, *8*, e428-436. https://doi.org/10.2741/1143

Gillain, S., Warzee, E., Lekeu, F., Wojtasik, V., Maquet, D., Croisier, J. L., . . . Petermans, J. (2009). The value of instrumental gait analysis in elderly healthy, MCI or Alzheimer's disease subjects and a comparison with other clinical tests used in single and dual-task conditions. *Annals of Physical and Rehabilitation Medicine*, *52*(6), 453–474. https://doi.org/10.1016/j.rehab.2008.10.004

Giné-Garriga, M., Roqué-Fíguls, M., Coll-Planas, L., Sitjà-Rabert, M., & Salvà, A. (2014). Physical exercise interventions for improving performance-based measures of physical function in community-dwelling, frail older adults: A systematic review and meta-analysis. *Archives of Physical Medicine and Rehabilitation*, *95*(4), 753-769.e3. https://doi.org/10.1016/j.apmr.2013.11.007

Gligoroska, J. P., & Manchevska, S. (2012). The effect of physical activity on cognition - physiological mechanisms. *Materia Socio-Medica*, *24*(3), 198–202. https://doi.org/10.5455/msm.2012.24.198-202

Gonçalves, A.-C., Cruz, J., Marques, A., Demain, S., & Samuel, D. (2018). Evaluating physical activity in dementia: A systematic review of outcomes to inform the development of a core outcome set. *Age and Ageing*, *47*(1), 34–41. https://doi.org/10.1093/ageing/afx135

Graessel, E., Viegas, R., Stemmer, R., Kuchly, B., Kornhuber, J., & Donath, C. (2009). The Erlangen Test of Activities of Daily Living: First results on reliability and validity of a short performance test to measure fundamental activities of daily living in dementia patients. *International Psychogeriatrics*, *21*(1), 103–112. https://doi.org/10.1017/S1041610208007710

Graham, J. E., Ostir, G. V., Kuo, Y.-F., Fisher, S. R., & Ottenbacher, K. J. (2008). Relationship between test methodology and mean velocity in timed walk tests: A review. *Archives of Physical Medicine and Rehabilitation*, *89*(5), 865–872. https://doi.org/10.1016/j.apmr.2007.11.029

Granacher, U., Bridenbaugh, S. A., Muehlbauer, T., Wehrle, A., & Kressig, R. W. (2011). Age-related effects on postural control under multi-task conditions. *Gerontology*, *57*(3), 247–255. https://doi.org/10.1159/000322196

Gras, L. Z., Kanaan, S. F., McDowd, J. M., Colgrove, Y. M., Burns, J. M., & Pohl, P. S. (2015). Balance and gait of adults with very mild Alzheimer disease. *Journal of Geriatric Physical Therapy (2001)*, *38*(1), 1–7. https://doi.org/10.1519/JPT.0000000000000020

Groot, C., Hooghiemstra, A. M., Raijmakers, P G H M, van Berckel, B N M, Scheltens, P., Scherder, E. J. A., . . . Ossenkoppele, R. (2016). The effect of physical activity on cognitive function in patients with dementia: A meta-analysis of randomized control trials. *Ageing Research Reviews*, *25*, 13–23. https://doi.org/10.1016/j.arr.2015.11.005

Guralnik, J. M., Ferrucci, L., Simonsick, E. M., Salive, M. E., & Wallace, R. B. (1995). Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. *The New England Journal of Medicine*, *332*(9), 556–561. https://doi.org/10.1056/NEJM199503023320902

Guralnik, J. M., Seeman, T. E., Tinetti, M. E., Nevitt, M. C., & Berkman, L. F. (1994). Validation and use of performance measures of functioning in a non-disabled older population: Macarthur studies of successful aging. *Aging (Milan, Italy)*, *6*(6), 410–419. https://doi.org/10.1007/BF03324272

Guralnik, J. M., Simonsick, E. M., Ferrucci, L., Glynn, R. J., Berkman, L. F., Blazer, D. G., . . . Wallace, R. B. (1994). A short physical performance battery assessing lower extremity function: Association with self-reported disability and prediction of mortality and nursing home admission. *Journal of Gerontology*, *49*(2), M85-94. https://doi.org/10.1093/geronj/49.2.M85

Haaland, K. Y. (2015). Disorders of higher cortical function. In L. E. Davis & S. Pirio Richardson (Eds.), *Fundamentals of neurologic disease* (pp. 133–145). New York: Springer. https://doi.org/10.1007/978-1-4939-2359-5_11

Hageman, P. A., & Thomas, V. S. (2002). Gait performance in dementia: The effects of a 6-week resistance training program in an adult day-care setting. *International Journal of Geriatric Psychiatry*, *17*(4), 329–334. https://doi.org/10.1002/gps.597

Hajek, A., Brettschneider, C., Posselt, T., Lange, C., Mamone, S., Wiese, B., . . . König, H.-H. (2016). Predictors of frailty in old age - results of a longitudinal study. *The Journal of Nutrition, Health & Aging*, *20*(9), 952–957. https://doi.org/10.1007/s12603-015-0634-5

Halsband, U., & Lange, R. K. (2006). Motor learning in man: A review of functional and clinical studies. *Journal of Physiology, Paris*, *99*(4-6), 414–424. https://doi.org/10.1016/j.jphysparis.2006.03.007

Harlein, J., Dassen, T., Halfens, R. J. G., & Heinze, C. (2009). Fall risk factors in older people with dementia or cognitive impairment: A systematic review. *Journal of Advanced Nursing*, *65*(5), 922–933. https://doi.org/10.1111/j.1365-2648.2008.04950.x

Hauer, K., Becker, C., Lindemann, U., & Beyer, N. (2006). Effectiveness of physical training on motor performance and fall prevention in cognitively impaired older persons: A systematic review. *American Journal of Physical Medicine & Rehabilitation*, *85*(10), 847–857. https://doi.org/10.1097/01.phm.0000228539.99682.32

Hauer, K., & Oster, P. (2008). Measuring functional performance in persons with dementia. *Journal of the American Geriatrics Society*, *56*(5), 949–950. https://doi.org/10.1111/j.1532-5415.2008.01649.x

Hauer, K., Schwenk, M., Zieschang, T., Essig, M., Becker, C., & Oster, P. (2012). Physical training improves motor performance in people with dementia: A randomized controlled trial. *Journal of the American Geriatrics Society*, *60*(1), 8–15. https://doi.org/10.1111/j.1532-5415.2011.03778.x

Hauer, K., Ullrich, P., Dutzi, I., Beurskens, R., Kern, S., Bauer, J. M., & Schwenk, M. (2017). Effects of standardized home training in patients with cognitive impairment following geriatric rehabilitation: A randomized controlled pilot study. *Gerontology*, *63*(6), 495–506. https://doi.org/10.1159/000478263

Hausdorff, J. M. (2007). Gait dynamics, fractals and falls: Finding meaning in the stride-to-stride fluctuations of human walking. *Human Movement Science*, *26*(4), 555–589. https://doi.org/10.1016/j.humov.2007.05.003

Hausdorff, J. M., & Buchman, A. S. (2013). What links gait speed and MCI with dementia? A fresh look at the association between motor and cognitive function. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *68*(4), 409–411. https://doi.org/10.1093/gerona/glt002

Hausdorff, J. M., Yogev, G., Springer, S., Simon, E. S., & Giladi, N. (2005). Walking is more like catching than tapping: Gait in the elderly as a complex cognitive task.

Experimental Brain Research, *164*(4), 541–548. https://doi.org/10.1007/s00221-005-2280-3

Hawk, C., Hyland, J. K., Rupert, R., Colonvega, M., & Hall, S. (2006). Assessment of balance and risk for falls in a sample of community-dwelling adults aged 65 and older. *Chiropractic & Osteopathy*, *14*(3). https://doi.org/10.1186/1746-1340-14-3

Henskens, M., Nauta, I. M., Drost, K. T., & Scherder, E. J. A. (2018). The effects of movement stimulation on activities of daily living performance and quality of life in nursing home residents with dementia: A randomized controlled trial. *Clinical Interventions in Aging*, *13*, 805–817. https://doi.org/10.2147/CIA.S160031

Hernández, S. S. S., Sandreschi, P. F., da Silva, F. C., Arancibia, B. A. V., da Silva, R., Gutierres, P. J. B., & Andrade, A. (2015). What are the benefits of exercise for Alzheimer's disease? A systematic review of the past 10 years. *Journal of Aging and Physical Activity*, *23*(4), 659–668. https://doi.org/10.1123/japa.2014-0180

Heyn, P. C., Johnson, K. E., & Kramer, A. F. (2008). Endurance and strength training outcomes on cognitively impaired and cognitively intact older adults: A meta-analysis. *The Journal of Nutrition, Health & Aging*, *12*(6), 401–409. https://doi.org/10.1007/BF02982674

Hill, K. D. (1996). A new test of dynamic standing balance for stroke patients: Reliability, validity and comparison with healthy elderly. *Physiotherapy Canada*, *48*(4), 257–262. https://doi.org/10.3138/ptc.48.4.257

Hogan, D. B., Bailey, P., Carswell, A., Clarke, B., Cohen, C., Forbes, D., . . . Thorpe, L. (2007). Management of mild to moderate Alzheimer's disease and dementia. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *3*(4), 355–384. https://doi.org/10.1016/j.jalz.2007.07.006

Hortobágyi, T., Lesinski, M., Gäbler, M., VanSwearingen, J. M., Malatesta, D., & Granacher, U. (2015). Effects of three types of exercise interventions on healthy old adults' gait speed: A systematic review and meta-analysis. *Sports Medicine (Auckland, N.Z.)*, *45*(12), 1627–1643. https://doi.org/10.1007/s40279-015-0371-2

Huang, S.-L., Hsieh, C.-L., Wu, R.-M., Tai, C.-H., Lin, C.-H., & Lu, W.-S. (2011). Minimal detectable change of the timed "up & go" test and the dynamic gait index in people with Parkinson disease. *Physical Therapy*, *91*(1), 114–121. https://doi.org/10.2522/ptj.20090126

Hüger, D., Zieschang, T., Schwenk, M., Oster, P., Becker, C., & Hauer, K. (2009). Designing studies on the effectiveness of physical training in patients with cognitive impairment. *Zeitschrift Für Gerontologie Und Geriatrie*, *42*(1), 11–19. https://doi.org/10.1007/s00391-008-0529-8

Hugo, J., & Ganguli, M. (2014). Dementia and cognitive impairment: Epidemiology, diagnosis, and treatment. *Clinics in Geriatric Medicine*, *30*(3), 421–442. https://doi.org/10.1016/j.cger.2014.04.001

Hüsing, B. (2010). Individualisierte Medizin - Potenziale und Handlungsbedarf [Individualised medicine - potentials and need for action]. *Zeitschrift fur Evidenz, Fortbildung und Qualität im Gesundheitswesen*, *104*(10), 727–731. https://doi.org/10.1016/j.zefq.2010.07.041

Hüsing, B., Hartig, J., Bührlen, B., Reiß, T., & Gaisser, S. (2008). Individualisierte Medizin und Gesundheitssystem. Retrieved from https://www.tab-beim-bundestag.de/de/pdf/publikationen/berichte/TAB-Arbeitsbericht-ab126.pdf

Ihl, R., & Weyer, G. (1993). *Die Alzheimer's Disease Assessment Scale (ADAS) [Alzheimer Disease Assessment Scale (ADAS)]*. Weinheim: Beltz Test.

Ijmker, T., & Lamoth, C. J. (2012). Gait and cognition: The relationship between gait stability and variability with executive function in persons with and without dementia. *Gait & Posture*, *35*(1), 126–130. https://doi.org/10.1016/j.gaitpost.2011.08.022

Jahn, K., Zwergal, A., & Schniepp, R. (2010). Gait disturbances in old age: Classification, diagnosis, and treatment from a neurological perspective. *Deutsches Ärzteblatt International*, *107*(17), 306-315. https://doi.org/10.3238/arztebl.2010.0306

Janssen, W. G. M., Bussmann, H. B. J., & Stam, H. J. (2002). Determinants of the sit-to-stand movement: A review. *Physical Therapy*, *82*(9), 866–879. https://doi.org/10.1093/ptj/82.9.866

Jebsen, R. H., Taylor, N., Trieschmann, R. B., Trotter, M. J., & Howard, L. A. (1969). An objective and standardized test of hand function. *Archives of Physical Medicine and Rehabilitation*, *50*(6), 311–319.

Jensen, C. S., Hasselbalch, S. G., Waldemar, G., & Simonsen, A. H. (2015). Biochemical markers of physical exercise on mild cognitive impairment and dementia: Systematic review and perspectives. *Frontiers in Neurology*, *6*, 187. https://doi.org/10.3389/fneur.2015.00187

Johansson, G., & Jarnlo, G.-B. (2009). Balance training in 70-year-old women. *Physiotherapy Theory and Practice*, *7*(2), 121–125. https://doi.org/10.3109/09593989109106962

Jones, C. J., Rikli, R. E., & Beam, W. C. (1999). A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Research Quarterly for Exercise and Sport*, *70*(2), 113–119. https://doi.org/10.1080/02701367.1999.10608028

Jones, C. J., Rikli, R. E., Max, J., & Noffal, G. (1998). The reliability and validity of a chair sit-and-reach test as a measure of hamstring flexibility in older adults. *Research Quarterly for Exercise and Sport*, *69*(4), 338–343. https://doi.org/10.1080/02701367.1998.10607708

Junge, T., Knudsen, H. K., & Kristensen, H. K. (2018). The effect of long-term, group-based physical, cognitive and social activities on physical performance in elderly, community-dwelling people with mild to moderate dementia. *Dementia (London, England).* Advance online publication. https://doi.org/10.1177/1471301218806376

Kampragkou, C., Iakovidis, P., Kampragkou, E., & Kellis, E. (2017). Effects of a 12-week aerobic exercise program combined with music therapy and memory exercises on cognitive and functional ability in people with middle type of Alzheimer's disease. *International Journal of Physiotherapy*, *4*(5), 262–268. https://doi.org/10.15621/ijphy/2017/v4i5/159420

Karantzoulis, S., & Galvin, J. E. (2011). Distinguishing Alzheimer's disease from other major forms of dementia. *Expert Review of Neurotherapeutics*, *11*(11), 1579–1591. https://doi.org/10.1586/ern.11.155

Karssemeijer, E. G. A., Aaronson, J. A., Bossers, W. J. R., Smits, T., Olde Rikkert, M. G. M., & Kessels, R. P. C. (2017). Positive effects of combined cognitive and physical exercise training on cognitive function in older adults with mild cognitive impairment or dementia: A meta-analysis. *Ageing Research Reviews*, *40*, 75–83. https://doi.org/10.1016/j.arr.2017.09.003

Kato-Narita, E. M., Nitrini, R., & Radanovic, M. (2011). Assessment of balance in mild and moderate stages of Alzheimer's disease: Implications on falls and functional capacity. *Arquivos De Neuro-Psiquiatria*, *69*(2A), 202–207. https://doi.org/10.1590/s0004-282x2011000200012

Kemoun, G., Thibaud, M., Roumagne, N., Carette, P., Albinet, C., Toussaint, L., . . . Dugue, B. (2010). Effects of a physical training programme on cognitive function and walking efficiency in elderly persons with dementia. *Dementia and Geriatric Cognitive Disorders*, *29*(2), 109–114. https://doi.org/10.1159/000272435

Kempler, D., & Goral, M. (2008). Language and dementia: Neuropsychological aspects. *Annual Review of Applied Linguistics*, *28*, 73–90. https://doi.org/10.1017/S0267190508080045

Kenigsberg, P.-A., Aquino, J.-P., Bérard, A., Gzil, F., Andrieu, S., Banerjee, S., . . . Robert, P. (2016). Dementia beyond 2025: Knowledge and uncertainties. *Dementia (London, England)*, *15*(1), 6–21. https://doi.org/10.1177/1471301215574785

Kennedy, G., Hardman, R. J., Macpherson, H., Scholey, A. B., & Pipingas, A. (2017). How does exercise reduce the rate of age-associated cognitive decline? A review of potential mechanisms. *Journal of Alzheimer's Disease: JAD*, *55*(1), 1–18. https://doi.org/10.3233/JAD-160665

Khachaturian, Z. S., Mesulam, M. M., Khachaturian, A. S., & Mohs, R. C. (2015). The special topics section of Alzheimer's & dementia. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *11*(11), 1261–1264. https://doi.org/10.1016/j.jalz.2015.10.002

Kido, T., Tabara, Y., Igase, M., Ochi, N., Uetani, E., Nagai, T., . . . Kohara, K. (2010). Postural instability is associated with brain atrophy and cognitive impairment in the elderly: The J-SHIPP study. *Dementia and Geriatric Cognitive Disorders*, *29*(5), 379–387. https://doi.org/10.1159/000255106

Kim, J.-S., Kang, M.-H., Moon, D.-C., & Oh, J.-S. (2017). Effect of pelvic tilt exercise using pressure-based visual biofeedback training on the gait parameter in elderly patients with Alzheimer's disease. *European Geriatric Medicine*, *8*(1), 30–36. https://doi.org/10.1016/j.eurger.2016.11.001

Kim, M.-J., Han, C.-W., Min, K.-Y., Cho, C.-Y., Lee, C.-W., Ogawa, Y., . . . Kohzuki, M. (2016). Physical exercise with multicomponent cognitive intervention for older adults with Alzheimer's disease: A 6-month randomized controlled trial. *Dementia and Geriatric Cognitive Disorders Extra*, *6*(2), 222–232. https://doi.org/10.1159/000446508

Kirova, A.-M., Bays, R. B., & Lagalwar, S. (2015). Working memory and executive function decline across normal aging, mild cognitive impairment, and Alzheimer's disease. *BioMed Research International.* Advance online publication. https://doi.org/10.1155/2015/748212

Knopman, D. S., Haeberlein, S. B., Carrillo, M. C., Hendrix, J. A., Kerchner, G., Margolin, R., . . . Siemers, E. (2018). The National Institute on Aging and the Alzheimer's Association Research Framework for Alzheimer's disease: Perspectives from the Research Roundtable. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *14*(4), 563–575. https://doi.org/10.1016/j.jalz.2018.03.002

Koo, T. K., & Li, M. Y. (2016). A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Kovács, E., Sztruhár Jónásné, I., Karóczi, C. K., Korpos, A., & Gondos, T. (2013). Effects of a multimodal exercise program on balance, functional mobility and fall risk in older adults with cognitive impairment: A randomized controlled single-blind study. *European Journal of Physical and Rehabilitation Medicine*, *49*(5), 639–648.

Kressig, R. W., & Beauchet, O. (2006). Guidelines for clinical applications of spatio-temporal gait analysis in older adults. *Aging Clinical and Experimental Research*, *18*(2), 174–176. https://doi.org/10.1007/BF03327437

Kuiack, S. L., Campbell, W. W., & Evans, W. J. (2004). A structured resistive training program improves muscle strength and power in elderly persons with dementia. *Activities, Adaptation & Aging*, *28*(1), 35–47. https://doi.org/10.1300/J016v28n01_03

Kumfor, F., Halliday, G. M., & Piguet, O. (2017). Clinical aspects of Alzheimer's disease. In P. Beart, M. Robinson, M. Rattray, & N. J. Maragakis (Eds.), *Neurodegenerative diseases* (pp. 31–53). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-57193-5_2

Lam, F. M. H., Huang, M.-Z., Liao, L.-R., Chung, R. C. K., Kwok, T. C. Y., & Pang, M. Y. C. (2018). Physical exercise improves strength, balance, mobility, and endurance in people with cognitive impairment and dementia: A systematic review. *Journal of Physiotherapy*, *64*(1), 4–15. https://doi.org/10.1016/j.jphys.2017.12.001

Lam, F. M. H., Liao, L.-R., Kwok, T. C. Y., & Pang, M. Y. C. (2018). Effects of adding whole-body vibration to routine day activity program on physical functioning in elderly with mild or moderate dementia: A randomized controlled trial. *International Journal of Geriatric Psychiatry*, *33*(1), 21–30. https://doi.org/10.1002/gps.4662

Lamb, S. E., Sheehan, B., Atherton, N., Nichols, V., Collins, H., Mistry, D., . . . Lall, R. (2018). Dementia And Physical Activity (DAPA) trial of moderate to high intensity exercise training for people with dementia: Randomised controlled trial. *BMJ (Clinical Research Ed.)*, *361.* https://doi.org/10.1136/bmj.k1675

Lamoth, C. J., van Deudekom, F. J., van Campen, J. P., Appels, B. A., Vries, O. J. de, & Pijnappels, M. (2011). Gait stability and variability measures show effects of impaired cognition and dual tasking in frail people. *Journal of Neuroengineering and Rehabilitation*, *8*(2). https://doi.org/10.1186/1743-0003-8-2

Latham, N. K., Bennett, D. A., Stretton, C. M., & Anderson, C. S. (2004). Systematic review of progressive resistance strength training in older adults. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *59*(1), 48–61. https://doi.org/10.1093/gerona/59.1.M48

Lautenschlager, N. T., Cox, K., & Cyarto, E. V. (2012). The influence of exercise on brain aging and dementia. *Biochimica Et Biophysica Acta*, *1822*(3), 474–481. https://doi.org/10.1016/j.bbadis.2011.07.010

Leandri, M., Cammisuli, S., Cammarata, S., Baratto, L., Campbell, J., Simonini, M., & Tabaton, M. (2009). Balance features in Alzheimer's disease and amnestic mild cognitive impairment. *Journal of Alzheimer's Disease: JAD*, *16*(1), 113–120. https://doi.org/10.3233/JAD-2009-0928

Lee, A. Y. (2011). Vascular dementia. *Chonnam Medical Journal*, *47*(2), 66–71. https://doi.org/10.4068/cmj.2011.47.2.66

Lee, H.-S. [Han-Suk], & Park, S.-W. (2017). The reliability of balance, gait, and muscle strength test for the elderly with dementia: A systematic review. *Journal of the Korean Society of Physical Medicine*, *12*(3), 49–58. https://doi.org/10.13066/kspm.2017.12.3.49

Lee, H.-S. [Han-Suk], Park, S.-W., & Chung, H. K. (2017). The Korean version of relative and absolute reliability of gait and balance assessment tools for patients with dementia in day care center and nursing home. *Journal of Physical Therapy Science*, *29*(11), 1934–1939. https://doi.org/10.1589/jpts.29.1934

Lehéricy, S., Bardinet, E., Tremblay, L., van de Moortele, P.-F., Pochon, J.-B., Dormont, D., . . . Ugurbil, K. (2006). Motor control in basal ganglia circuits using fMRI and brain atlas approaches. *Cerebral Cortex (New York, N.Y.: 1991)*, *16*(2), 149–161. https://doi.org/10.1093/cercor/bhi089

Lezak, M. D. (2012). *Neuropsychological assessment* (5th ed.). Oxford, New York: Oxford University Press.

Liberati, A., Altman, D. G., Tetzlaff, J. M., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., . . . Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Medicine*, *6*(7), e1000100. https://doi.org/10.1371/journal.pmed.1000100

Lin, S.-Y., & Lewis, F. M. (2015). Dementia friendly, dementia capable, and dementia positive: Concepts to prepare for the future. *The Gerontologist*, *55*(2), 237–244. https://doi.org/10.1093/geront/gnu122

Lindeboom, J., & Weinstein, H. (2004). Neuropsychology of cognitive ageing, minimal cognitive impairment, Alzheimer's disease, and vascular cognitive impairment. *European Journal of Pharmacology*, *490*(1-3), 83–86. https://doi.org/10.1016/j.ejphar.2004.02.046

Linn, B. S., Linn, M. W., & Gurel, L. (1968). Cumulative illness rating scale. *Journal of the American Geriatrics Society*, *16*(5), 622–626. https://doi.org/10.1111/j.1532-5415.1968.tb02103.x

Littbrand, H., Stenvall, M., & Rosendahl, E. (2011). Applicability and effects of physical exercise on physical and cognitive functions and activities of daily living among people with dementia: A systematic review. *American Journal of Physical Medicine & Rehabilitation*, *90*(6), 495–518. https://doi.org/10.1097/PHM.0b013e318214de26

Liu, C.-j., & Latham, N. K. (2009). Progressive resistance strength training for improving physical function in older adults. *The Cochrane Database of Systematic Reviews*. (3), CD002759. https://doi.org/10.1002/14651858.CD002759.pub2

Liu, C.-j., Shiroy, D. M., Jones, L. Y., & Clark, D. O. (2014). Systematic review of functional training on muscle strength, physical functioning, and activities of daily living in older adults. *European Review of Aging and Physical Activity*, *11*(2), 95–106. https://doi.org/10.1007/s11556-014-0144-1

Liu-Ambrose, T., & Donaldson, M. G. (2009). Exercise and cognition in older adults: Is there a role for resistance training programmes? *British Journal of Sports Medicine*, *43*(1), 25–27. https://doi.org/10.1136/bjsm.2008.055616

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., . . . Mukadam, N. (2017). Dementia prevention, intervention, and care.

Lancet (London, England), *390*(10113), 2673–2734. https://doi.org/10.1016/S0140-6736(17)31363-6

Lord, S. R., Menz, H. B., & Tiedemann, A. (2003). A physiological profile approach to falls risk assessment and prevention. *Physical Therapy*, *83*(3), 237–252. https://doi.org/10.1093/ptj/83.3.237

Lord, S. R., Murray, S. M., Chapman, K., Munro, B., & Tiedemann, A. (2002). Sit-to-stand performance depends on sensation, speed, balance, and psychological status in addition to strength in older people. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *57*(8), M539-543. https://doi.org/10.1093/gerona/57.8.M539

Lübke, N., Meinck, M., & Renteln-Kruse, W. von (2004). The Barthel Index in geriatrics. A context analysis for the Hamburg Classification Manual [Der Barthel-Index in der Geriatrie. Eine Kontextanalyse zum Hamburger Einstufungsmanual]. *Zeitschrift für Gerontologie und Geriatrie*, *37*(4), 316–326. https://doi.org/10.1007/s00391-004-0233-2

Lundin-Olsson, L., Nyberg, L., & Gustafson, Y. (1998). Attention, frailty, and falls: The effect of a manual task on basic mobility. *Journal of the American Geriatrics Society*, *46*(6), 758–761. https://doi.org/10.1111/j.1532-5415.1998.tb03813.x

Luttenberger, K., Schmiedeberg, A., & Grassel, E. (2012). Activities of daily living in dementia: Revalidation of the E-ADL Test and suggestions for further development. *BMC Psychiatry*, *12*(208). https://doi.org/10.1186/1471-244X-12-208

Mahlknecht, P., Kiechl, S., Bloem, B. R., Willeit, J., Scherfler, C., Gasperi, A., . . . Seppi, K. (2013). Prevalence and burden of gait disorders in elderly men and women aged 60-97 years: A population-based study. *PloS One*, *8*(7), e69627. https://doi.org/10.1371/journal.pone.0069627

Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: The Barthel index. *Maryland State Medical Journal*, *14*, 61–65.

Mancini, M., & Horak, F. B. (2010). The relevance of clinical balance assessment tools to differentiate balance deficits. *European Journal of Physical and Rehabilitation Medicine*, *46*(2), 239–248.

Manckoundia, P., Mourey, F., Pfitzenmeyer, P., & Papaxanthis, C. (2006). Comparison of motor strategies in sit-to-stand and back-to-sit motions between healthy and Alzheimer's disease elderly subjects. *Neuroscience*, *137*(2), 385–392. https://doi.org/10.1016/j.neuroscience.2005.08.079

Manckoundia, P., Taroux, M., Kubicki, A., & Mourey, F. (2014). Impact of ambulatory physiotherapy on motor abilities of elderly subjects with Alzheimer's disease. *Geriatrics & Gerontology International*, *14*(1), 167–175. https://doi.org/10.1111/ggi.12075

Maquet, D., Lekeu, F., Warzee, E., Gillain, S., Wojtasik, V., Salmon, E., . . . Croisier, J. L. (2010). Gait analysis in elderly adult patients with mild cognitive impairment and patients with mild Alzheimer's disease: Simple versus dual task: A preliminary report. *Clinical Physiology and Functional Imaging*, *30*(1), 51–56. https://doi.org/10.1111/j.1475-097X.2009.00903.x

Martyr, A., & Clare, L. (2012). Executive function and activities of daily living in Alzheimer's disease: A correlational meta-analysis. *Dementia and Geriatric Cognitive Disorders*, *33*(2-3), 189–203. https://doi.org/10.1159/000338233

Mathias, S., Nayak, U. S., & Isaacs, B. (1986). Balance in elderly patients: The "get-up and go" test. *Archives of Physical Medicine and Rehabilitation*, *67*(6), 387–389.

Matsunaga, S., Kishi, T., & Iwata, N. (2015a). Combination therapy with cholinesterase inhibitors and memantine for Alzheimer's disease: A systematic review and meta-analysis. *The International Journal of Neuropsychopharmacology*, *18*(5). https://doi.org/10.1093/ijnp/pyu115

Matsunaga, S., Kishi, T., & Iwata, N. (2015b). Memantine monotherapy for Alzheimer's disease: A systematic review and meta-analysis. *PloS One*, *10*(4), e0123289. https://doi.org/10.1371/journal.pone.0123289

Mazoteras Muñoz, V., van Abellan Kan, G., Cantet, C., Cortes, F., Ousset, P.-J., Rolland, Y. M., & Vellas, B. (2010). Gait and balance impairments in Alzheimer disease patients. *Alzheimer Disease and Associated Disorders*, *24*(1), 79–84. https://doi.org/10.1097/WAD.0b013e3181c78a20

Mc Ardle, R., Morris, R., Wilson, J., Galna, B., Thomas, A. J., & Rochester, L. (2017). What can quantitative gait analysis tell us about dementia and its subtypes? A structured review. *Journal of Alzheimer's Disease: JAD*, *60*(4), 1295–1312. https://doi.org/10.3233/JAD-170541

McCaffrey, R., Park, J., Newman, D., & Hagen, D. (2014). The effect of chair yoga in older adults with moderate and severe Alzheimer's disease. *Research in Gerontological Nursing*, *7*(4), 171–177. https://doi.org/10.3928/19404921-20140218-01

McCarthy, E. K., Horvat, M. A., Holtsberg, P. A., & Wisenbaker, J. M. (2004). Repeated chair stands as a measure of lower limb strength in sexagenarian women. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *59*(11), 1207–1212. https://doi.org/10.1093/gerona/59.11.1207

McGough, E. L., Lin, S.-Y., Belza, B., Becofsky, K. M., Jones, D. L., Liu, M., . . . Logsdon, R. G. (2019). A scoping review of physical performance outcome measures used in exercise interventions for older adults with Alzheimer disease and related dementias. *Journal of Geriatric Physical Therapy (2001)*, *42*(1), 28–47. https://doi.org/10.1519/JPT.0000000000000159

McGough, E. L., Logsdon, R. G., Kelly, V. E., & Teri, L. (2013). Functional mobility limitations and falls in assisted living residents with dementia: Physical performance assessment and quantitative gait analysis. *Journal of Geriatric Physical Therapy (2001)*, *36*(2), 78–86. https://doi.org/10.1519/JPT.0b013e318268de7f

McKhann, G., Drachman, D., Folstein, M. F., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, *34*(7), 939–944. https://doi.org/10.1212/wnl.34.7.939

Mendes-Santos, L. C., Mograbi, D., Spenciere, B., & Charchat-Fichman, H. (2015). Specific algorithm method of scoring the Clock Drawing Test applied in cognitively normal elderly. *Dementia & Neuropsychologia*, *9*(2), 128–135. https://doi.org/10.1590/1980-57642015DN92000007

Merory, J. R., Wittwer, J. E., Rowe, C. C., & Webster, K. E. (2007). Quantitative gait analysis in patients with dementia with Lewy bodies and Alzheimer's disease. *Gait & Posture*, *26*(3), 414–419. https://doi.org/10.1016/j.gaitpost.2006.10.006

Mian, O. S., Baltzopoulos, V., Minetti, A. E., & Narici, M. V. (2007). The impact of physical training on locomotor function in older people. *Sports Medicine (Auckland, N.Z.)*, *37*(8), 683–701. https://doi.org/10.2165/00007256-200737080-00003

Mijnarends, D. M., Meijers, J. M. M., Halfens, R. J. G., ter Borg, S., Luiking, Y. C., Verlaan, S., . . . Schols, J. M. G. A. (2013). Validity and reliability of tools to measure muscle mass, strength, and physical performance in community-dwelling older people: A systematic review. *Journal of the American Medical Directors Association*, *14*(3), 170–178. https://doi.org/10.1016/j.jamda.2012.10.009

Mitnitski, A. B., Graham, J. E., Mogilner, A. J., & Rockwood, K. (1999). The rate of decline in function in Alzheimer's disease and other dementias. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *54*(2), M65-69. https://doi.org/10.1093/gerona/54.2.M65

Miu, D., Szeto, S., & Mak, Y. (2008). A randomized controlled trial on the effect of exercise on physical, cognitive, and affective function in dementia subjects. *Asian Journal of Gerontology and Geriatrics*, *3*(1), 8–16.

Moe-Nilssen, R. (1998). Test-retest reliability of trunk accelerometry during standing and walking. *Archives of Physical Medicine and Rehabilitation*, *79*(11), 1377–1385. https://doi.org/10.1016/S0003-9993(98)90231-3

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., . . . Altman, D. G. (2010). Consort 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ (Clinical Research Ed.)*, *340*, c869. https://doi.org/10.1136/bmj.c869

Moher, D., Liberati, A., Tetzlaff, J. M., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mokkink, L. B., Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C., & Terwee, C. B. (2018). COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs): User manual. Retrieved from https://www.cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf

Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., . . . Vet, H. C. W. de (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, *63*(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Mokkink, L. B., Vet, H. C. W. de, Prinsen, C. A., Patrick, D. L., Alonso, J., Bouter, L. M., & Terwee, C. B. (2018). Cosmin Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *27*(5), 1171–1179. https://doi.org/10.1007/s11136-017-1765-4

Montero-Odasso, M., Verghese, J., Beauchet, O., & Hausdorff, J. M. (2012). Gait and cognition: A complementary approach to understanding brain function and the risk of falling. *Journal of the American Geriatrics Society*, *60*(11), 2127–2136. https://doi.org/10.1111/j.1532-5415.2012.04209.x

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., . . . Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*(9), 1159–1165. https://doi.org/10.1212/WNL.39.9.1159

Morris, J. C., Mohs, R. C., Rogers, H., Fillenbaum, G., & Heyman, A. (1988). Consortium to establish a registry for Alzheimer's disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. *Psychopharmacology Bulletin*, *24*(4), 641–652.

Morris, R., Lord, S. [Susan], Bunce, J., Burn, D. J., & Rochester, L. (2016). Gait and cognition: Mapping the global and discrete relationships in ageing and neurodegenerative disease. *Neuroscience and Biobehavioral Reviews*, *64*, 326–345. https://doi.org/10.1016/j.neubiorev.2016.02.012

Mosa, A. S. M., Yoo, I., & Sheets, L. (2012). A systematic review of healthcare applications for smartphones. *BMC Medical Informatics and Decision Making*, *12*(67). https://doi.org/10.1186/1472-6947-12-67

Muir, S. W., Speechley, M., Wells, J., Borrie, M., Gopaul, K., & Montero-Odasso, M. (2012). Gait assessment in mild cognitive impairment and Alzheimer's disease: The effect of dual-task challenges across the cognitive spectrum. *Gait & Posture*, *35*(1), 96–100. https://doi.org/10.1016/j.gaitpost.2011.08.014

Muir-Hunter, S. W., Graham, L., & Montero-Odasso, M. (2015). Reliability of the Berg Balance Scale as a clinical measure of balance in community-dwelling older adults with mild to moderate Alzheimer disease: A pilot study. *Physiotherapy Canada*, *67*(3), 255–262. https://doi.org/10.3138/ptc.2014-32

Nadkarni, N. K., Mawji, E., McIlroy, W. E., & Black, S. E. (2009). Spatial and temporal gait parameters in Alzheimer's disease and aging. *Gait & Posture*, *30*(4), 452–454. https://doi.org/10.1016/j.gaitpost.2009.07.003

Nadkarni, N. K., McIlroy, W. E., Mawji, E., & Black, S. E. (2009). Gait and subcortical hyperintensities in mild Alzheimer's disease and aging. *Dementia and Geriatric Cognitive Disorders*, *28*(4), 295–301. https://doi.org/10.1159/000245158

Nakamura, T., Meguro, K., Yamazaki, H., Okuzumi, H., Tanaka, A., Horikawa, A., . . . Sasaki, H. (1997). Postural and gait disturbance correlated with decreased frontal cerebral blood flow in Alzheimer disease. *Alzheimer Disease and Associated Disorders*, *11*(3), 132–139. https://doi.org/10.1097/00002093-199709000-00005

Nakayama, N., Suzuki, M., & Hamaguchi, T. (2019). Relationship between knee extension strength and gait styles in patients with dementia. *Medicine*, *98*(12), e14958. https://doi.org/10.1097/MD.0000000000014958

Nelson, P. T., Alafuzoff, I., Bigio, E. H., Bouras, C., Braak, H., Cairns, N. J., . . . Beach, T. G. (2012). Correlation of Alzheimer disease neuropathologic changes with cognitive status: A review of the literature. *Journal of Neuropathology and Experimental Neurology*, *71*(5), 362–381. https://doi.org/10.1097/NEN.0b013e31825018f7

Netz, Y., Axelrad, S., & Argov, E. (2007). Group physical activity for demented older adults feasibility and effectiveness. *Clinical Rehabilitation*, *21*(11), 977–986. https://doi.org/10.1177/0269215507078318

Netz, Y., Ayalon, M., Dunsky, A., & Alexander, N. B. (2004). 'the multiple-sit-to-stand' field test for older adults: What does it measure? *Gerontology*, *50*(3), 121–126. https://doi.org/10.1159/000076769

Öhman, H., Savikko, N., Strandberg, T. E., & Pitkälä, K. H. (2014). Effect of physical exercise on cognitive performance in older adults with mild cognitive impairment or dementia: A systematic review. *Dementia and Geriatric Cognitive Disorders*, *38*(5-6), 347–365. https://doi.org/10.1159/000365388

Olsen, C. F., & Bergland, A. (2017). "Reliability of the Norwegian version of the short physical performance battery in older people with and without dementia". *BMC Geriatrics*, *17*(124). https://doi.org/10.1186/s12877-017-0514-4

Orange, J. B., Molloy, D. W., Lever, J. A., Darzins, P., & Ganesan, C. R. (1994). Alzheimer's disease. Physician-patient communication. *Canadian Family Physician Medecin De Famille Canadien*, *40*(20), 1160–1168. https://doi.org/10.1001/jama.1994.03520200075039

Orcioli-Silva, D., Barbieri, F. A., Simieli, L., Santos, P. C. R. d., Beretta, V. S., Coelho, F. G. d. M., . . . Gobbi, L. T. B. (2018). A program of physical activity improves gait impairment in people with Alzheimer's disease. *Motriz: Revista De Educação Física*, *24*(1). https://doi.org/10.1590/s1980-6574201800010009

Padala, K. P., Padala, P. R., Lensing, S. Y., Dennis, R. A., Bopp, M. M., Roberson, P. K., & Sullivan, D. H. (2017). Home-based exercise program improves balance and fear of falling in community-dwelling older adults with mild Alzheimer's disease: A pilot study. *Journal of Alzheimer's Disease: JAD*, *59*(2), 565–574. https://doi.org/10.3233/JAD-170120

Padala, K. P., Padala, P. R., Malloy, T. R., Geske, J. A., Dubbert, P. M., Dennis, R. A., . . . Sullivan, D. H. (2012). Wii-fit for improving gait and balance in an assisted living facility: A pilot study. *Journal of Aging Research.* Advance online publication. https://doi.org/10.1155/2012/597573

Parihar, R., Mahoney, J. R., & Verghese, J. (2013). Relationship of gait and cognition in the elderly. *Current Translational Geriatrics and Experimental Gerontology Reports*, *2*(3). https://doi.org/10.1007/s13670-013-0052-7

Pavasini, R., Guralnik, J. M., Brown, J. C., Di Bari, M., Cesari, M., Landi, F., . . . Campo, G. (2016). Short Physical Performance Battery and all-cause mortality: Systematic review and meta-analysis. *BMC Medicine*, *14*(1), 215. https://doi.org/10.1186/s12916-016-0763-7

Pedrinolla, A., Venturelli, M., Fonte, C., Munari, D., Benetti, M. V., Rudi, D., . . . Schena, F. (2018). Exercise training on locomotion in patients with Alzheimer's disease: A feasibility study. *Journal of Alzheimer's Disease: JAD*, *61*(4), 1599–1609. https://doi.org/10.3233/JAD-170625

Perrochon, A., Tchalla, A. E., Bonis, J., Perucaud, F., & Mandigout, S. (2015). Effects of a multicomponent exercise program on spatiotemporal gait parameters, risk of falling and physical activity in dementia patients. *Dementia and Geriatric Cognitive Disorders Extra*, *5*(3), 350–360. https://doi.org/10.1159/000435772

Perry, R. J., & Hodges, J. R. (1999). Attention and executive deficits in Alzheimer's disease. A critical review. *Brain: A Journal of Neurology*, *122*(3), 383–404. https://doi.org/10.1093/brain/122.3.383

Pettersson, A. F., Engardt, M., & Wahlund, L.-O. (2002). Activity level and balance in subjects with mild Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *13*(4), 213–216. https://doi.org/10.1159/000057699

Pettersson, A. F., Olsson, E., & Wahlund, L.-O. (2005). Motor function in subjects with mild cognitive impairment and early Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *19*(5-6), 299–304. https://doi.org/10.1159/000084555

Phillips, C. D., Chu, C. W., Morris, J. N., & Hawes, C. (1993). Effects of cognitive impairment on the reliability of geriatric assessments in nursing homes. *Journal of the American Geriatrics Society*, *41*(2), 136–142. https://doi.org/10.1111/j.1532-5415.1993.tb02047.x

Pistollato, F., Ohayon, E. L., Lam, A., Langley, G. R., Novak, T. J., Pamies, D., . . . Chandrasekera, P. C. (2016). Alzheimer disease research in the 21st century: Past and current failures, new perspectives and funding priorities. *Oncotarget*, *7*(26), 38999–39016. https://doi.org/10.18632/oncotarget.9175

Pitkälä, K. H., Pöysti, M. M., Laakkonen, M.-L., Tilvis, R. S., Savikko, N., Kautiainen, H., & Strandberg, T. E. (2013). Effects of the Finnish Alzheimer disease exercise trial (FINALEX): A randomized controlled trial. *JAMA Internal Medicine*, *173*(10), 894–901. https://doi.org/10.1001/jamainternmed.2013.359

Pitkälä, K. H., Savikko, N., Pöysti, M. M., Strandberg, T. E., & Laakkonen, M.-L. (2013). Efficacy of physical exercise intervention on mobility and physical functioning in older people with dementia: A systematic review. *Experimental Gerontology*, *48*(1), 85–93. https://doi.org/10.1016/j.exger.2012.08.008

Podsiadlo, D., & Richardson, S. (1991). The timed "Up & Go": A test of basic functional mobility for frail elderly persons. *Journal of the American Geriatrics Society*, *39*(2), 142–148. https://doi.org/10.1111/j.1532-5415.1991.tb01616.x

Pomeroy, V. (1990). Development of an ADL oriented assessment-of-mobility scale suitable for use with elderly people with dementia. *Physiotherapy*, *76*(8), 446–448. https://doi.org/10.1016/S0031-9406(10)62968-9

Pomeroy, V., Warren, C. M., Honeycombe, C., Briggs, R. S., Wilkinson, D. G., Pickering, R. M., & Steiner, A. (1999). Mobility and dementia: Is physiotherapy treatment during respite care effective? *International Journal of Geriatric Psychiatry*, *14*(5), 389–397. https://doi.org/10.1002/(SICI)1099-1166(199905)14:5<389::AID-GPS933>3.0.CO;2-8

Portney, L. G., & Watkins, M. P. (2015). *Foundations of clinical research: Applications to practice* (3rd ed. [rev.]). Upper Saddle River, N.J.: Pearson/Prentice Hall.

Possin, K. L. (2010). Visual spatial cognition in neurodegenerative disease. *Neurocase*, *16*(6), 466–487. https://doi.org/10.1080/13554791003730600

Potter, R., Ellard, D., Rees, K., & Thorogood, M. (2011). A systematic review of the effects of physical activity on physical functioning, quality of life and depression in older people with dementia. *International Journal of Geriatric Psychiatry*, *26*(10), 1000–1011. https://doi.org/10.1002/gps.2641

Prince, M. J., Guerchet, M., & Prina, M. (2015). The epidemiology and impact of dementia: Current state and future trends. Retrieved from http://www.who.int/mental_health/neurology/dementia/dementia_thematicbrief_epidemiology.pdf

Prince, M. J., Guerchet, M., Prina, M., & Alzheimer's Disease International (2013). Policy brief: The global impact of dementia 2013–2050. Retrieved from https://www.alz.co.uk/research/GlobalImpactDementia2013.pdf

Prinsen, C. A., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., Vet, H. C. W. de, & Terwee, C. B. (2018). Cosmin guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *27*(5), 1147–1157. https://doi.org/10.1007/s11136-018-1798-3

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills*, *8*(3), 271–276. https://doi.org/10.2466/pms.1958.8.3.271

Reitan, R. M. (1992). *Trail Making Test: Manual for administration and scoring.*

Reith, W., & Muhl-Benninghaus, R. (2015). Differenzialdiagnose demenzieller Erkrankungen [Differential diagnostics of dementia type diseases]. *Der Radiologe*, *55*(5), 378–385. https://doi.org/10.1007/s00117-014-2799-z

Reuben, D. B., & Siu, A. L. (1990). An objective measure of physical function of elderly outpatients. The Physical Performance Test. *Journal of the American Geriatrics Society*, *38*(10), 1105–1112. https://doi.org/10.1111/j.1532-5415.1990.tb01373.x

Ries, J. D., Echternach, J. L., Nof, L., & Gagnon Blodgett, M. (2009). Test-retest reliability and minimal detectable change scores for the timed "up & go" test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Physical Therapy*, *89*(6), 569–579. https://doi.org/10.2522/ptj.20080258

Ries, J. D., Hutson, J., Maralit, L. A., & Brown, M. B. (2015). Group balance training specifically designed for individuals with Alzheimer disease: Impact on Berg Balance Scale, Timed Up and Go, Gait Speed, and Mini-Mental Status Examination. *Journal of Geriatric Physical Therapy (2001)*, *38*(4), 183–193. https://doi.org/10.1519/JPT.0000000000000030

Rikli, R. E., & Jones, C. J. (2006). *Senior fitness test manual - Updated Edition*. Champaign, IL: Human Kinetics.

Ringhof, S., & Stein, T. (2018). Biomechanical assessment of dynamic balance: Specificity of different balance tests. *Human Movement Science*, *58*, 140–147. https://doi.org/10.1016/j.humov.2018.02.004

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron*, *31*(6), 889–901. https://doi.org/10.1016/S0896-6273(01)00423-8

Roach, K. E., Tappen, R. M., Kirk-Sanchez, N., Williams, C. L., & Loewenstein, D. (2011). A randomized controlled trial of an activity specific exercise program for individuals with Alzheimer disease in long-term care settings. *Journal of Geriatric Physical Therapy (2001)*, *34*(2), 50–56. https://doi.org/10.1519/JPT.0b013e31820aab9c

Roach, K. E., & van Dillen, L. R. (1988). Development of an Acute Care Index of functional status for patients with neurologic impairment. *Physical Therapy*, *68*(7), 1102–1108. https://doi.org/10.1093/ptj/68.7.1102

Rockwood, K. (2002). Vascular cognitive impairment and vascular dementia. *Journal of the Neurological Sciences*, *203-204*, 23–27. https://doi.org/10.1016/S0022-510X(02)00255-1

Rockwood, K., Awalt, E., Carver, D., & MacKnight, C. (2000). Feasibility and measurement properties of the functional reach and the timed up and go tests in the Canadian study of health and aging. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *55*(2), M70-73. https://doi.org/10.1093/gerona/55.2.M70

Rockwood, K., Rockwood, M. R. H., Andrew, M. K., & Mitnitski, A. B. (2008). Reliability of the hierarchical assessment of balance and mobility in frail older adults. *Journal of the American Geriatrics Society*, *56*(7), 1213–1217. https://doi.org/10.1111/j.1532-5415.2008.01773.x

Rolland, Y. M., Cesari, M., Miller, M. E., Penninx, B. W., Atkinson, H. H., & Pahor, M. (2004). Reliability of the 400-m usual-pace walk test as an assessment of mobility limitation in older adults. *Journal of the American Geriatrics Society*, *52*(6), 972–976. https://doi.org/10.1111/j.1532-5415.2004.52267.x

Rolland, Y. M., Pillard, F., Klapouszczak, A., Reynish, E., Thomas, D. R., Andrieu, S., . . . Vellas, B. (2007). Exercise program for nursing home residents with Alzheimer's disease: A 1-year randomized, controlled trial. *Journal of the American Geriatrics Society*, *55*(2), 158–165. https://doi.org/10.1111/j.1532-5415.2007.01035.x

Rosano, C., Kuller, L. H., Chung, H. [Hyoju], Arnold, A. M., Longstreth, W. T., & Newman, A. B. (2005). Subclinical brain magnetic resonance imaging abnormalities predict physical functional decline in high-functioning older adults. *Journal of the American Geriatrics Society*, *53*(4), 649–654. https://doi.org/10.1111/j.1532-5415.2005.53214.x

Rossiter-Fornoff, J. E., Wolf, S. L., Wolfson, L. I., & Buchner, D. M. (1995). A cross-sectional validation study of the FICSIT common data base static balance measures. Frailty and Injuries: Cooperative Studies of Intervention Techniques. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *50*(6), M291-297. https://doi.org/10.1093/gerona/50A.6.M291

Rucco, R., Agosti, V., Jacini, F., Sorrentino, P., Varriale, P., Stefano, M. de, . . . Sorrentino, G. (2017). Spatio-temporal and kinematic gait analysis in patients with Frontotemporal dementia and Alzheimer's disease through 3D motion capture. *Gait & Posture*, *52*, 312–317. https://doi.org/10.1016/j.gaitpost.2016.12.021

Ruhe, A., Fejer, R., & Walker, B. (2010). The test-retest reliability of centre of pressure measures in bipedal static task conditions - a systematic review of the literature. *Gait & Posture*, *32*(4), 436–445. https://doi.org/10.1016/j.gaitpost.2010.09.012

Rydwik, E., Frändin, K., & Akner, G. (2004). Effects of physical training on physical performance in institutionalised elderly patients (70+) with multiple diagnoses. *Age and Ageing*, *33*(1), 13–23. https://doi.org/10.1093/ageing/afh001

Saghaei, M., & Saghaei, S. (2011). Implementation of an open-source customizable minimization program for allocation of patients to parallel groups in clinical trials. *Journal of Biomedical Science and Engineering*, *4*(11), 734–739. https://doi.org/10.4236/jbise.2011.411090

Salardini, A. (2019). An overview of primary dementias as clinicopathological entities. *Seminars in Neurology*, *39*(2), 153–166. https://doi.org/10.1055/s-0039-1683445

Santana-Sosa, E., Barriopedro, M. I., López-Mojares, L. M., Pérez, M., & Lucia, A. (2008). Exercise training is beneficial for Alzheimer's patients. *International Journal of Sports Medicine*, *29*(10), 845–850. https://doi.org/10.1055/s-2008-1038432

Scharpf, A., Servay, S., & Woll, A. (2013). Auswirkungen von körperlicher Aktivität auf demenzielle Erkrankungen. *Sportwissenschaft*, *43*(3), 166–180. https://doi.org/10.1007/s12662-013-0295-7

Scherder, E. J. A., Eggermont, L. H., Swaab, D., van Heuvelen, M. J. G., Kamsma, Y., Greef, M. de, . . . Mulder, T. (2007). Gait in ageing and associated dementias; its relationship with cognition. *Neuroscience and Biobehavioral Reviews*, *31*(4), 485–497. https://doi.org/10.1016/j.neubiorev.2006.11.007

Schneider, J. A., Li, J.-l. [Jia-liang], Li, Y. [Yan], Wilson, R. S., Kordower, J. H., & Bennett, D. A. (2006). Substantia nigra tangles are related to gait impairment in older persons. *Annals of Neurology*, *59*(1), 166–173. https://doi.org/10.1002/ana.20723

Schulz, K. F., Altman, D. G., & Moher, D. (2010). Consort 2010 statement: Updated guidelines for reporting parallel group randomized trials. *Annals of Internal Medicine*, *152*(11), 726–732. https://doi.org/10.7326/0003-4819-152-11-201006010-00232

Schulze, J., van den Bussche, H., Kaduszkiewicz, H., Koller, D., & Hoffmann, F. (2015). Institutionalization in incident dementia cases in comparison to age- and sex- matched controls: A 5-year follow-up from Germany. *Social Psychiatry and Psychiatric Epidemiology*, *50*(1), 143–151. https://doi.org/10.1007/s00127-014-0911-3

Schünemann, H., Brożek, J., Guyatt, G., & Oxman, A. (2013). GRADE handbook for grading quality of evidence and strength of recommendations. Updated October 2013. Retrieved from https://gdt.gradepro.org/app/handbook/handbook.html

Schwenk, M., Dutzi, I., Englert, S., Micol, W., Najafi, B., Mohler, J., & Hauer, K. (2014). An intensive exercise program improves motor performances in patients with dementia: Translational model of geriatric rehabilitation. *Journal of Alzheimer's Disease: JAD*, *39*(3), 487–498. https://doi.org/10.3233/JAD-130470

Schwenk, M., Gogulla, S., Englert, S., Czempik, A., & Hauer, K. (2012). Test-retest reliability and minimal detectable change of repeated sit-to-stand analysis using one body fixed sensor in geriatric patients. *Physiological Measurement*, *33*(11), 1931–1946. https://doi.org/10.1088/0967-3334/33/11/1931

Schwenk, M., Hauer, K., Zieschang, T., Englert, S., Mohler, J., & Najafi, B. (2014). Sensor-derived physical activity parameters can predict future falls in people with dementia. *Gerontology*, *60*(6), 483–492. https://doi.org/10.1159/000363136

Schwenk, M., Schmidt, M. [Marita], Pfisterer, M., Oster, P., & Hauer, K. (2011). Rollator use adversely impacts on assessment of gait and mobility during geriatric rehabilitation. *Journal of Rehabilitation Medicine*, *43*(5), 424–429. https://doi.org/10.2340/16501977-0791

Schwenk, M., Zieschang, T., Englert, S., Grewal, G., Najafi, B., & Hauer, K. (2014). Improvements in gait characteristics after intensive resistance and functional training in people with dementia: A randomised controlled trial. *BMC Geriatrics*, *14*(73). https://doi.org/10.1186/1471-2318-14-73

Schwenk, M., Zieschang, T., Oster, P., & Hauer, K. (2010). Dual-task performances can be improved in patients with dementia: A randomized controlled trial. *Neurology*, *74*(24), 1961–1968. https://doi.org/10.1212/WNL.0b013e3181e39696

Shah, H., Albanese, E., Duggan, C., Rudan, I., Langa, K. M., Carrillo, M. C., . . . Dua, T. (2016). Research priorities to reduce the global burden of dementia by 2025. *The Lancet. Neurology*, *15*(12), 1285–1294. https://doi.org/10.1016/S1474-4422(16)30235-6

Sheridan, P. L., & Hausdorff, J. M. (2007). The role of higher-level cognitive function in gait: Executive dysfunction contributes to fall risk in Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*, *24*(2), 125–137. https://doi.org/10.1159/000105126

Sheridan, P. L., Solomont, J., Kowall, N., & Hausdorff, J. M. (2003). Influence of executive function on locomotor function: Divided attention increases gait variability in Alzheimer's disease. *Journal of the American Geriatrics Society*, *51*(11), 1633–1637. https://doi.org/10.1046/j.1532-5415.2003.51516.x

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Shulman, K. I., Shedletsky, R., & Silver, I. L. (1986). The challenge of time: Clock-drawing and cognitive function in the elderly. *International Journal of Geriatric Psychiatry*, *1*(2), 135–140. https://doi.org/10.1002/gps.930010209

Shumway-Cook, A., Brauer, S., & Woollacott, M. (2000). Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Physical Therapy*, *80*(9), 896–903. https://doi.org/10.1093/ptj/80.9.896

Smidt, N., van der Windt, D. A., Assendelft, W. J., Mourits, A. J., Devillé, W. L., Winter, A. F. de, & Bouter, L. M. (2002). Interobserver reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis. *Archives of Physical Medicine and Rehabilitation*, *83*(8), 1145–1150. https://doi.org/10.1053/apmr.2002.33728

Sobol, N. A., Hoffmann, K., Frederiksen, K. S., Vogel, A., Vestergaard, K., Brændgaard, H., . . . Beyer, N. (2016). Effect of aerobic exercise on physical performance in patients with Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, *12*(12), 1207–1215. https://doi.org/10.1016/j.jalz.2016.05.004

Sosa-Ortiz, A. L., Acosta-Castillo, I., & Prince, M. J. (2012). Epidemiology of dementias and Alzheimer's disease. *Archives of Medical Research*, *43*(8), 600–608. https://doi.org/10.1016/j.arcmed.2012.11.003

Souto Barreto, P. de, Cesari, M., Denormandie, P., Armaingaud, D., Vellas, B., & Rolland, Y. M. (2017). Exercise or social intervention for nursing home residents with dementia: A pilot randomized, controlled trial. *Journal of the American Geriatrics Society*, *65*(9), E123-E129. https://doi.org/10.1111/jgs.14947

Steinberg, M., Leoutsakos, J.-M. S., Podewils, L. J., & Lyketsos, C. G. (2009). Evaluation of a home-based exercise program in the treatment of Alzheimer's disease: The Maximizing Independence in Dementia (MIND) study. *International Journal of Geriatric Psychiatry*, *24*(7), 680–685. https://doi.org/10.1002/gps.2175

Sterke, C. S., van Beeck, E. F., Looman, C. W. N., Kressig, R. W., & van der Cammen, Tischa J M (2012). An electronic walkway can predict short-term fall risk in nursing home residents with dementia. *Gait & Posture*, *36*(1), 95–101. https://doi.org/10.1016/j.gaitpost.2012.01.012

Stewart, D. A., Burns, J. M., Dunn, S. G., & Roberts, M. A. (2016). The two-minute walking test: A sensitive index of mobility in the rehabilitation of elderly patients. *Clinical Rehabilitation*, *4*(4), 273–276. https://doi.org/10.1177/026921559000400404

Stopford, C. L., Thompson, J. C., Neary, D., Richardson, A. M. T., & Snowden, J. S. (2012). Working memory, attention, and executive function in Alzheimer's disease and frontotemporal dementia. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *48*(4), 429–446. https://doi.org/10.1016/j.cortex.2010.12.002

Suttanon, P., Hill, K. D., Dodd, K. J., & Said, C. M. (2011). Retest reliability of balance and mobility measurements in people with mild to moderate Alzheimer's disease. *International Psychogeriatrics*, *23*(7), 1152–1159. https://doi.org/10.1017/S1041610211000639

Suttanon, P., Hill, K. D., Said, C. M., & Dodd, K. J. (2010). Can balance exercise programmes improve balance and related physical performance measures in people with dementia? A systematic review. *European Review of Aging and Physical Activity*, *7*(1), 13–25. https://doi.org/10.1007/s11556-010-0055-8

Suttanon, P., Hill, K. D., Said, C. M., LoGiudice, D., Lautenschlager, N. T., & Dodd, K. J. (2012). Balance and mobility dysfunction and falls risk in older people with mild to moderate Alzheimer disease. *American Journal of Physical Medicine & Rehabilitation*, *91*(1), 12–23. https://doi.org/10.1097/PHM.0b013e31823caeea

Suttanon, P., Hill, K. D., Said, C. M., Williams, S. B., Byrne, K. N., LoGiudice, D., . . . Dodd, K. J. (2013). Feasibility, safety and preliminary evidence of the effectiveness of a home-based exercise programme for older people with Alzheimer's disease: A pilot randomized controlled trial. *Clinical Rehabilitation*, *27*(5), 427–438. https://doi.org/10.1177/0269215512460877

Suzuki, M., Kirimoto, H., Inamura, A., Yagi, M., Omori, Y., & Yamada, S. (2012). The relationship between knee extension strength and lower extremity functions in nursing home residents with dementia. *Disability and Rehabilitation*, *34*(3), 202–209. https://doi.org/10.3109/09638288.2011.593678

Suzuki, M., Yamada, S., Inamura, A., Omori, Y., Kirimoto, H., Sugimura, S., & Miyamoto, M. (2009). Reliability and validity of measurements of knee extension strength obtained from nursing home residents with dementia. *American Journal of Physical Medicine & Rehabilitation*, *88*(11), 924–933. https://doi.org/10.1097/PHM.0b013e3181ae1003

Tangen, G. G., Engedal, K., Bergland, A., Moger, T. A., & Mengshoel, A. M. (2014). Relationships between balance and cognition in patients with subjective cognitive impairment, mild cognitive impairment, and Alzheimer disease. *Physical Therapy*, *94*(8), 1123–1134. https://doi.org/10.2522/ptj.20130298

Tang-Wai, D. F., & Graham, N. L. (2008). Assessment of language function in dementia. *Geriatrics & Ageing*, *11*(2), 103–110.

Tappen, R. M., Roach, K. E., Applegate, E. B., & Stowell, P. (2000). Effect of a combined walking and conversation intervention on functional mobility of nursing home residents with Alzheimer disease. *Alzheimer Disease and Associated Disorders*, *14*(4), 196–201. https://doi.org/10.1097/00002093-200010000-00002

Tappen, R. M., Roach, K. E., Buchner, D. M., Barry, C., & Edelstein, J. (1997). Reliability of physical performance measures in nursing home residents with Alzheimer's

disease. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *52*(1), M52-55. https://doi.org/10.1093/gerona/52A.1.M52

Tay, L., Lim, W. S., Chan, M., Ali, N., & Chong, M. S. (2016). A combined cognitive stimulation and physical exercise programme (MINDVital) in early dementia: Differential effects on single- and dual-task gait performance. *Gerontology*, *62*(6), 604–610. https://doi.org/10.1159/000444084

Telenius, E. W., Engedal, K., & Bergland, A. (2013). Physical performance and quality of life of nursing-home residents with mild and moderate dementia. *International Journal of Environmental Research and Public Health*, *10*(12), 6672–6686. https://doi.org/10.3390/ijerph10126672

Telenius, E. W., Engedal, K., & Bergland, A. (2015a). Effect of a high-intensity exercise program on physical function and mental health in nursing home residents with dementia: An assessor blinded randomized controlled trial. *PloS One*, *10*(5), e0126102. https://doi.org/10.1371/journal.pone.0126102

Telenius, E. W., Engedal, K., & Bergland, A. (2015b). Inter-rater reliability of the Berg Balance Scale, 30 s chair stand test and 6 m walking test, and construct validity of the Berg Balance Scale in nursing home residents with mild-to-moderate dementia. *Bmj Open*, *5*(9), e008321. https://doi.org/10.1136/bmjopen-2015-008321

Thalheimer, W., & Cook, S. (2002). How to calculate effect sizes from published research: A simplified methodology. Retrieved from https://www.bwgriffin.com/gsu/courses/edur9131/content/Effect_Sizes_pdf5.pdf

Theill, N., Martin, M., Schumacher, V., Bridenbaugh, S. A., & Kressig, R. W. (2011). Simultaneously measuring gait and cognitive performance in cognitively healthy and cognitively impaired older adults: The Basel motor-cognition dual-task paradigm. *Journal of the American Geriatrics Society*, *59*(6), 1012–1018. https://doi.org/10.1111/j.1532-5415.2011.03429.x

Thom, J. M., & Clare, L. (2011). Rationale for combined exercise and cognition-focused interventions to improve functional independence in people with dementia. *Gerontology*, *57*(3), 265–275. https://doi.org/10.1159/000322198

Thomas, V. S., & Hageman, P. A. (1999). A preliminary study on the reliability of physical performance measures in older day-care center clients with dementia. *International Psychogeriatrics*, *14*(1), 17–23. https://doi.org/10.1017/S1041610202008244

Thomas, V. S., & Hageman, P. A. (2003). Can neuromuscular strength and function in people with dementia be rehabilitated using resistance-exercise training? Results from a preliminary intervention study. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *58*(8), 746–751. https://doi.org/10.1093/gerona/58.8.m746

Thomas, V. S., Vandenberg, E. V., & Potter, J. F. (2002). Non-neurological factors are implicated in impairments in gait and mobility among patients in a clinical dementia referral population. *International Journal of Geriatric Psychiatry*, *17*(2), 128–133. https://doi.org/10.1002/gps.547

Thurm, F., Scharpf, A., Liebermann, N., Kolassa, S., Elbert, T., Lüchtenberg, D., . . . Kolassa, I.-T. (2011). Improvement of Cognitive Function after Physical Movement Training in Institutionalized Very Frail Older Adults with Dementia. *GeroPsych*, *24*(4), 197–208. https://doi.org/10.1024/1662-9647/a000048

Tian, Q., Chastan, N., Bair, W.-N., Resnick, S. M., Ferrucci, L., & Studenski, S. A. (2017). The brain map of gait variability in aging, cognitive impairment and dementia - A systematic review. *Neuroscience and Biobehavioral Reviews*, *74*, 149–162. https://doi.org/10.1016/j.neubiorev.2017.01.020

Tiedemann, A., Sherrington, C., & Lord, S. R. (2005). Physiological and psychological predictors of walking speed in older community-dwelling people. *Gerontology*, *51*(6), 390–395. https://doi.org/10.1159/000088703

Tinetti, M. E. (1986). Performance-Oriented Assessment of Mobility problems in elderly patients. *Journal of the American Geriatrics Society*, *34*(2), 119–126. https://doi.org/10.1111/j.1532-5415.1986.tb05480.x

Toots, A., Littbrand, H., Holmberg, H., Nordström, P., Lundin-Olsson, L., Gustafson, Y., & Rosendahl, E. (2017). Walking aids moderate exercise effects on gait speed in people with dementia: A randomized controlled trial. *Journal of the American Medical Directors Association*, *18*(3), 227–233. https://doi.org/10.1016/j.jamda.2016.09.003

Toots, A., Littbrand, H., Lindelöf, N., Wiklund, R., Holmberg, H., Nordström, P., . . . Rosendahl, E. (2016). Effects of a high-intensity functional exercise program on dependence in activities of daily living and balance in older adults with dementia. *Journal of the American Geriatrics Society*, *64*(1), 55–64. https://doi.org/10.1111/jgs.13880

Toulotte, C., Fabre, C., Dangremont, B., Lensel, G., & Thévenon, A. (2003). Effects of physical training on the physical capacity of frail, demented patients with a history of falling: A randomised controlled trial. *Age and Ageing*, *32*(1), 67–73. https://doi.org/10.1093/ageing/32.1.67

Trautwein, S., Barisch-Fritz, B., Scharpf, A., Bossers, W. J. R., Meinzer, M., Steib, S., . . . Woll, A. (2019). Recommendations for assessing motor performance in individuals with dementia: Suggestions of an expert panel - a qualitative approach. *European Review of Aging and Physical Activity*, *16*(5). https://doi.org/10.1186/s11556-019-0212-7

Trautwein, S., Barisch-Fritz, B., Scharpf, A., Ringhof, S., Stein, T., Krell-Roesch, J., & Woll, A. (submitted). Effects of a 16-Week multimodal exercise program on gait performance in individuals with dementia: A multicenter randomized controlled trial.

Trautwein, S., Hadzic, A., Barisch-Fritz, B., Maurus, P., & Woll, A. (in prep.). Effectiveness of physical activity on temporal and spatial gait parameters in individuals with dementia: A systematic review.

Trautwein, S., Maurus, P., Barisch-Fritz, B., Hadzic, A., & Woll, A. (2019). Recommended motor assessments based on psychometric properties in individuals with dementia: A systematic review. *European Review of Aging and Physical Activity*, *16*(20). https://doi.org/10.1186/s11556-019-0228-z

Trautwein, S., Scharpf, A., Barisch-Fritz, B., Niermann, C., & Woll, A. (2017). Effectiveness of a 16-week multimodal exercise program on individuals with dementia: Study protocol for a multicenter randomized controlled trial. *JMIR Research Protocols*, *6*(3), e35. https://doi.org/10.2196/resprot.6792

U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, & U.S. Department of Health and Human Services

FDA Center for Devices and Radiological Health (2006). Guidance for industry: Patient-reported outcome measures: Use in medical product development to support labeling claims: Draft guidance. *Health and Quality of Life Outcomes*, *4*(79). https://doi.org/10.1186/1477-7525-4-79

Valkanova, V., & Ebmeier, K. P. (2017). What can gait tell us about dementia? Review of epidemiological and neuropsychological evidence. *Gait & Posture*, *53*, 215–223. https://doi.org/10.1016/j.gaitpost.2017.01.024

Van Iersel, M. B., Benraad, C. E. M., & Olde Rikkert, M. G. M. (2007). Validity and reliability of quantitative gait analysis in geriatric patients with and without dementia. *Journal of the American Geriatrics Society*, *55*(4), 632–634. https://doi.org/10.1111/j.1532-5415.2007.01130.x

Van Iersel, M. B., Hoefsloot, W., Munneke, M., Bloem, B. R., & Olde Rikkert, M. G. M. (2004). Systematic review of quantitative clinical gait analysis in patients with dementia. *Zeitschrift Für Gerontologie Und Geriatrie*, *37*(1), 27–32. https://doi.org/10.1007/s00391-004-0176-7

Van Iersel, M. B., Munneke, M., Esselink, R. A. J., Benraad, C. E. M., & Olde Rikkert, M. G. M. (2008). Gait velocity and the Timed-Up-and-Go test were sensitive to changes in mobility in frail elderly patients. *Journal of Clinical Epidemiology*, *61*(2), 186–191. https://doi.org/10.1016/j.jclinepi.2007.04.016

Vasunilashorn, S., Coppin, A. K., Patel, K. V., Lauretani, F., Ferrucci, L., Bandinelli, S., & Guralnik, J. M. (2009). Use of the Short Physical Performance Battery score to predict loss of ability to walk 400 meters: Analysis from the InCHIANTI study. *The Journals of Gerontology Series a: Biological Sciences and Medical Sciences*, *64*(2), 223–229. https://doi.org/10.1093/gerona/gln022

Vellas, B., Rubenstein, L. Z., Ousset, P.-J., Faisant, C., Kostek, V., Nourhashemi, F., . . . Albarede, J. L. (1997). One-leg standing balance and functional status in a population of 512 community-living elderly persons. *Aging Clinical and Experimental Research*, *9*(1-2), 95–98. https://doi.org/10.1007/BF03340133

Ventola, C. L. (2014). Mobile devices and apps for health care professionals: Uses and benefits. *P & T: A Peer-Reviewed Journal for Formulary Management*, *39*(5), 356–364.

Venturelli, M., Scarsini, R., & Schena, F. (2011). Six-month walking program changes cognitive and ADL performance in patients with Alzheimer. *American Journal of Alzheimer's Disease and Other Dementias*, *26*(5), 381–388. https://doi.org/10.1177/1533317511418956

Verghese, J., Annweiler, C., Ayers, E., Barzilai, N., Beauchet, O., Bennett, D. A., . . . Wang, C. (2014). Motoric cognitive risk syndrome: Multicountry prevalence and dementia risk. *Neurology*, *83*(8), 718–726. https://doi.org/10.1212/WNL.0000000000000717

Verghese, J., Wang, C., Lipton, R. B., Holtzer, R., & Xue, X. (2007). Quantitative gait dysfunction and risk of cognitive decline and dementia. *Journal of Neurology, Neurosurgery, and Psychiatry*, *78*(9), 929–935. https://doi.org/10.1136/jnnp.2006.106914

Verkerke, G. J., Lemmink, K. A. P. M., Slagers, A. J., Westhoff, M. H., van Riet, G. A. J., & Rakhorst, G. (2003). Precision, comfort and mechanical performance of the Quadriso-tester, a quadriceps force measuring device. *Medical and*

Biological Engineering and Computing, *41*(3), 283–289. https://doi.org/10.1007/BF02348432

Vermeulen, J., Neyens, J. C. L., van Rossum, E., Spreeuwenberg, M. D., & Witte, L. P. de (2011). Predicting ADL disability in community-dwelling elderly people using physical frailty indicators: A systematic review. *BMC Geriatrics*, *11*(33). https://doi.org/10.1186/1471-2318-11-33

Veronese, N., Bolzetta, F., Toffanello, E. D., Zambon, S., Rui, M. de, Perissinotto, E., . . . Manzato, E. (2014). Association between Short Physical Performance Battery and falls in older people: The Progetto Veneto Anziani Study. *Rejuvenation Research*, *17*(3), 276–284. https://doi.org/10.1089/rej.2013.1491

Versijpt, J. (2014). Effectiveness and cost-effectiveness of the pharmacological treatment of Alzheimer's disease and vascular dementia. *Journal of Alzheimer's Disease: JAD*, *42*(Suppl 3), S19-25. https://doi.org/10.3233/JAD-132639

Visser, H. (1983). Gait and balance in senile dementia of Alzheimer's type. *Age and Ageing*, *12*(4), 296–301. https://doi.org/10.1093/ageing/12.4.296

Voelcker-Rehage, C., & Niemann, C. (2013). Structural and functional brain changes related to different types of physical activity across the life span. *Neuroscience and Biobehavioral Reviews*, *37*, 2268–2295. https://doi.org/10.1016/j.neubiorev.2013.01.028

Vreugdenhil, A., Cannell, J., Davies, A., & Razay, G. (2012). A community-based exercise programme to improve functional ability in people with Alzheimer's disease: A randomized controlled trial. *Scandinavian Journal of Caring Sciences*, *26*(1), 12–19. https://doi.org/10.1111/j.1471-6712.2011.00895.x

Waldemar, G., Dubois, B., Emre, M., Georges, J., McKeith, I. G., Rossor, M., . . . Winblad, B. (2007). Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: Efns guideline. *European Journal of Neurology*, *14*(1), e1-26. https://doi.org/10.1111/j.1468-1331.2006.01605.x

Walker, L., McAleese, K. E., Erskine, D., & Attems, J. (2019). Neurodegenerative diseases and ageing. In J. R. Harris & V. I. Korolchuk (Eds.), *Biochemistry and cell biology of ageing: Part II clinical science* (pp. 75–106). Singapore: Springer. https://doi.org/10.1007/978-981-13-3681-2_4

Webster, K. E., Merory, J. R., & Wittwer, J. E. (2006). Gait variability in community dwelling adults with Alzheimer disease. *Alzheimer Disease and Associated Disorders*, *20*(1), 37–40. https://doi.org/10.1097/01.wad.0000201849.75578.de

Weintraub, S., Wicklund, A. H., & Salmon, D. P. (2012). The neuropsychological profile of Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine*, *2*(4). https://doi.org/10.1101/cshperspect.a006171

Wennie Huang, W.-N., Perera, S., VanSwearingen, J. M., & Studenski, S. A. (2010). Performance measures predict onset of activity of daily living difficulty in community-dwelling older adults. *Journal of the American Geriatrics Society*, *58*(5), 844–852. https://doi.org/10.1111/j.1532-5415.2010.02820.x

Werner, C., Wiloth, S., Lemke, N. C., Kronbach, F., & Hauer, K. (2018). Development and validation of a novel motor-cognitive assessment strategy of compensatory sit-to-stand maneuvers in people with dementia. *Journal of Geriatric Physical Therapy (2001)*, *41*(3), 143–154. https://doi.org/10.1519/JPT.0000000000000116

Werner, C., Wiloth, S., Lemke, N. C., Kronbach, F., Jansen, C.-P., Oster, P., . . . Hauer, K. (2017). People with dementia can learn compensatory movement maneuvers for the sit-to-stand task: A randomized controlled trial. *Journal of Alzheimer's Disease: JAD*, *60*(1), 107–120. https://doi.org/10.3233/JAD-170258

Wesson, J., Clemson, L., Brodaty, H., Lord, S. R., Taylor, M. E., Gitlin, L. N., & Close, J. (2013). A feasibility study and pilot randomised trial of a tailored prevention program to reduce falls in older people with mild dementia. *BMC Geriatrics*, *13*(89). https://doi.org/10.1186/1471-2318-13-89

Wilde, N. J., Strauss, E., & Tulsky, D. S. (2004). Memory span on the Wechsler Scales. *Journal of Clinical and Experimental Neuropsychology*, *26*(4), 539–549. https://doi.org/10.1080/13803390490496605

Wilmore, J. H., & Costill, D. L. (2004). *Physiology of sport and exercise* (3. ed.). Champaign, Ill.: Human Kinetics.

Wiloth, S., Lemke, N. C., Werner, C., & Hauer, K. (2016). Validation of a computerized, game-based assessment strategy to measure training effects on motor-cognitive functions in people with dementia. *JMIR Serious Games*, *4*(2), e12. https://doi.org/10.2196/games.5696

Wiloth, S., Werner, C., Lemke, N. C., Bauer, J. M., & Hauer, K. (2018). Motor-cognitive effects of a computerized game-based training method in people with dementia: A randomized controlled trial. *Aging & Mental Health*, *22*(9), 1124–1135. https://doi.org/10.1080/13607863.2017.1348472

Wittwer, J. E., Webster, K. E., Andrews, P. T., & Menz, H. B. (2008). Test-retest reliability of spatial and temporal gait parameters of people with Alzheimer's disease. *Gait & Posture*, *28*(3), 392–396. https://doi.org/10.1016/j.gaitpost.2008.01.007

Wittwer, J. E., Webster, K. E., & Hill, K. D. (2013). Reproducibility of gait variability measures in people with Alzheimer's disease. *Gait & Posture*, *38*(3), 507–510. https://doi.org/10.1016/j.gaitpost.2013.01.021

Wittwer, J. E., Webster, K. E., & Menz, H. B. (2010). A longitudinal study of measures of walking in people with Alzheimer's disease. *Gait & Posture*, *32*(1), 113–117. https://doi.org/10.1016/j.gaitpost.2010.04.001

World Health Organization (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva. Retrieved from http://www.who.int/classifications/icd/en/bluebook.pdf?ua=1

World Health Organization (2012). *Dementia: A public health priority*. Geneva, London: World Health Organization; Alzheimer's Disease International. Retrieved from https://extranet.who.int/agefriendlyworld/wp-content/uploads/2014/06/WHO-Dementia-English.pdf

World Health Organization (2016). International statistical classification of diseases and related health problems 10th revision (ICD-10). Retrieved from https://icd.who.int/browse10/2016/en#/F00

Yogev-Seligmann, G., Hausdorff, J. M., & Giladi, N. (2008). The role of executive function and attention in gait. *Movement Disorders: Official Journal of the Movement Disorder Society*, *23*(3), 329-342. https://doi.org/10.1002/mds.21720

Yoon, J. E., Lee, S. M., Lim, H. S., Kim, T. H., Jeon, J. K., & Mun, M. H. (2013). The effects of cognitive activity combined with active extremity exercise on balance,

walking activity, memory level and quality of life of an older adult sample with dementia. *Journal of Physical Therapy Science*, *25*(12), 1601–1604. https://doi.org/10.1589/jpts.25.1601

Yümin, E. T., Şimşek, T. T., Sertel, M., Öztürk, A., & Yümin, M. (2011). The effect of functional mobility and balance on health-related quality of life (HRQoL) among elderly people living at home and those living in nursing home. *Archives of Gerontology and Geriatrics*, *52*(3), e180-184. https://doi.org/10.1016/j.archger.2010.10.027

Zijlstra, W., & Aminian, K. (2007). Mobility assessment in older people: New possibilities and challenges. *European Journal of Ageing*, *4*(1), 3–12. https://doi.org/10.1007/s10433-007-0041-9

Appendix

Additional files manuscript II

Additional file 1

Search term first search

Details of the search strategy used in Pubmed are provided below. This search was modified as appropriate for other databases. All fields were searched for MeSH and free search terms:

1. dementia
2. “Alzheimer disease”
3. “Vascular dementia”
4. or/1-3
5. “physical activity”
6. “motor activity”
7. exercise
8. “physical training”
9. training
10. or/5-9
11. “functional performance”
12. fitness
13. “physical fitness”
14. balance
15. equilibrium
16. gait
17. mobility
18. strength
19. flexibility
20. endurance
21. or/11-20
22. 4 and 10 and 21

Additional file 2

*Formulas for calculating time*group interaction effect sizes*

$$d = \sqrt{F \cdot \left(\frac{n_{IG} + n_{CG}}{n_{IG} \cdot n_{CG}}\right) \cdot \left(\frac{n_{IG} + n_{CG}}{n_{IG} + n_{CG} - 2}\right)}$$

$$d = t \cdot \sqrt{\left(\frac{n_{IG} + n_{CG}}{n_{IG} \cdot n_{CG}}\right) \cdot \left(\frac{n_{IG} + n_{CG}}{n_{IG} + n_{CG} - 2}\right)}$$

$$d = \frac{mean_{DIFF\ IG} - mean_{DIFF\ CG}}{\sqrt{\frac{(n_{IG} - 1) \cdot SD^2_{DIFF\ IG} + (n_{CG} - 1) \cdot SD^2_{DIFF\ CG}}{n_{IG} + n_{CG}}}}$$

with d=Cohen’s d effect size, F=F statistic, $mean_{DIFF\ CG}$=baseline-post mean difference in control group, $mean_{DIFF\ IG}$=baseline-post mean difference in intervention group, n_{CG}=number of subjects in control group, n_{IG}= number of subjects in intervention group, $SD^2_{DIFF\ CG}$=standard deviation of baseline-post difference in control group, $SD^2_{DIFF\ IG}$= standard deviation of baseline-post difference in intervention group, t=t statistic (Thalheimer & Cook, 2002).

Additional file 3

Search term main search

Details of the search strategy used in Pubmed are provided below. This search was modified as appropriate for other databases. Title and abstract were searched for MeSH and free search terms:

1. dementia
2. Alzheimer*
3. “vascular dementia”
4. “frontotemporal disease”
5. or/1-4
6. validity
7. valid
8. “content validity”
9. “structural validity”
10. “criterion validity”
11. reliability
12. reliable
13. consistency
14. “Cronbach’s alpha”
15. reproducibility
16. repeatability
17. “intra-rater”
18. “intra rater”
19. “inter-rater”
20. “inter rater”
21. “relative reliability”
22. correlation
23. kappa
24. “intra class correlation”
25. “intra class”
26. “intra-class”
27. ICC
28. “limits of agreement”
29. LOA
30. “absolute reliability”
31. “standard error of measurement”
32. SEM
33. “minimal detectable change”
34. MDC
35. “smallest detectable change”
36. SDC
37. or/6-36
38. “near-tandem test”
39. “single leg stance”
40. “SLS”
72. “TUGT”
73. “timed get-up-and-go test”
74. “TGUG”
75. “modified timed up and go”
76. “TUG mod”
77. “get up and go test”
78. “get-up-and-go test”
79. “timed up and go test with a secondary cognitive”
80. “timed up and go test with a secondary motor”
81. “6-meter walk test”
82. “six-meter walking test”
83. “6-meter walking speed”
84. “10-m walk test”
85. “walking speed over 10 m”
86. “8-ft walk test”
87. “Timed 8-foot walk”
88. “4-m usual gait speed”
89. “gait analysis”
90. “gait performance”
91. “Bessou locometer”
92. “Southampton Assessment of Mobility”
93. “Southampton Mobility Assessment”
94. “Hierarchical Assessment of Balance and Mobility”
95. “HABAM”
96. “Sit-to-Stand”
97. “STS”
98. “Chair sit to stand test”
99. “CST”
100. “chair rise test”
101. “Timed Chair Stands”
102. “5-time-sit-to-stand test”
103. “5-chair-stand”
104. “30-second chair stand test”
105. “30-s chair stand test”
106. “30-second sit-to-stand test”
107. “stair-climbing performance”
108.” “arm curl test”
109. “handgrip”
110. “dynamometer
111. “one-repetition maximum”

41. “one-leg balance test”
42. “One Leg Standing Balance Test”
43. “OLST”
44. “frailty and injuries cooperative studies of intervention techniques – subtest 4”
45. “FICSIT-4”
46. “Posturography platform”
47. “Wii Balance Board”
48. “NeuroCom Balance Master”
49. “functional reach test”
50. “functional reach”
51. “FR”
52. “Hill Step Test”
53. “Step test”
54. “figure of eight test”
55. “Groningen meander walking test”
56. “Berg Balance Scale”
57. “BERG”
58. “BBS”
59. “modified Berg Balance Scale”
60. “m-BBS”
61. “Performance Oriented Mobility Assessment”
62. “POMA”
63. “Performance Oriented Motor Assessment”
64. “Tinetti’s Performance Oriented Motor Assessment”
65. “Tinetti scale”
66. “Tinetti Test”
67. “TT”
68. “Tinetti balance assessment”
69. “Tinetti Balance Evaluation Test”
70. “timed up and go”
71. “TUG”
112. “1-RM”
113. “leg press”
114. “Physical therapy assessment”
115. “2-min walk test”
116. “Two Minute Walking Test”
117. “6-minute walk test”
118. “6-minute walking test”
119. “6WT”
120. “The 6-Minute Walk”
121. “Modified 6-Minute Walk”
122. “400-m walk test”
123. “3-speed walking test”
124. “6-min Astrand Cycle Ergometer test”
125. “Ergometric test and rest electrocardiogram”
126. “ECG”
127. “pedal power”
128. “Chair sit and reach”
129. “Short Physical Performance Battery”
130. “SPPB”
131. “Physical Performance Test”
132. “PPT”
133. “7-item Physical Performance Test”
134. “PPT-7”
135. “Erlangen-ADL test”
136. “E-ADL”
137. “Senior Fitness test”
138. “Jebsen hand function test”
139. “JHFT”
140. “Jebsen Total Time”
141. “JTT”
142. “Physiological Profile Assessment”
143. “PPA”
144. or/38-143
145. 5 and 37 and 144

Additional file 4

Formulas for calculating minimal detectable change at 95 % confidence interval

$$MDC_{95\%} = \frac{MDC_{95}}{mean} \cdot 100$$

$$MDC_{95\%} = \frac{SEM \cdot 1.96 \cdot \sqrt{2}}{mean} \cdot 100$$

with MDC_{95}=minimal detectable change at 95 % confidence interval, mean=mean of all available scores for an assessment, and SEM=standard error of measurement (Portney & Watkins, 2015; Schwenk et al., 2012)

Additional file 5

Study characteristics first search

***Table A.** Study characteristics first search

Reference	Study de-sign	Sample size	Setting	Group	Sample characteristics				Intervention
					Gender [% female]	Age [years] Mean (SD)	MMSE Mean (SD)	Aetiology of dementia	
Aguiar et al. (2014)	RCT	n_r=40 n_a=34	n.r.	CG	71 %	74.7 (7.4)	20.8 (4.0)	AD	Rivastigmine transdermal patch
				IG	77 %	78.6 (8.4)	20.1 (4.5)		Rivastigmine transdermal patch + exercise (aerobic, flexibility, strength, balance; 2x/week, 40min)
				All	74 %	n.r.	n.r., range: n.r.		6 months
Arcoverde et al. (2014)	RCT	n_r=20 n_a=20	Outpa-tient unit	CG	50 %	79.0 (n.r.)	19.9 (3.4)	AD, MD	No intervention
				IG	60 %	78.5 (n.r.)	20.4 (2.7)		Treadmill walking (2x/week, 30min)
				All	n.r.	n.r.	n.r., CI_{95}: 15-25		3 months
Bossers et al. (2015)	RCT	n_r=123 n_a=109	Care facility	CG	69 %	85.4 (5.0)	15.9 (4.2)	AD, VD, MD, unknown	Social visits (4x/week, 30min)
				IG1	78 %	85.7 (5.1)	15.8 (4.3)		Strength + walking (4x/week, 30min)
				IG2	78 %	85.4 (5.4)	15.2 (4.8)		Walking (4x/week, 30min)
				All	n.r.	85.5 (5.1)	n.r., range: 9-23		9 weeks
Bossers et al. (2016)	RCT	n_r=118 n_a=105	Care facility	CG	69 %	85.7 (4.8)	15.9 (4.3)	AD, VD, MD, unknown	Social visits (4x/week, 30min)
				IG1	77 %	85.7 (5.2)	15.9 (4.4)		Strength + walking (4x/week, 30min)
				IG2	77 %	85.5 (5.4)	15.3 (4.8)		Walking (4x/week, 30min)
				All	n.r.	85.6 (5.1)	n.r., range: 9-23		9 weeks

Burgener et al. (2008)	RCT	n_r=43 n_a=33	n.r.	CG	47 %	76.0 (8.1)	22.9 (5.2)	AD, VD, MD, LBD, FLD	Attention-control educational program
				IG	46 %	77.9 (7.9)	24.8 (3.5)		Taiji exercise (3x/week, 60min), cognitive-behavioural therapy, support group (both bi-weekly, 90min)
				All	47 %	77.0 (n.r.)	n.r., range: n.r.		20 weeks
Cancela et al. (2016)	RCT	n_r=189 n_a=189/114*	Care facility	CG	81 %	82.9 (7.4)	n.r.	n.r.	Non-physical distractive recreational activities
				IG	44 %	80.6 (8.3)	n.r.		Cycling (daily, min. 15min)
				All	67 %	n.r.	n.r., range: n.r.		15 months
Christofoletti et al. (2008)	RCT	n_r=54 n_a=41	Long-term psychiatric institution	CG	70 %	79.4 (2.0)	14.6 (1.2)	MD	No motor intervention
				IG1	65 %	70.0 (1.8)	18.7 (1.7)		Physiotherapy, occupational therapy, physical education (5x/week, 120min)
				IG2	71 %	72.9 (2.3)	12.7 (2.1)		Physiotherapy (3x/week, 60min)
				All	69 %	74.3 (1.4)	n.r., range: n.r.		6 months
Cott et al. (2002)	RCT	n_r=86 n_a=74	Care facility	CG	42 %	79.8 (8.3)	6.31 (7.46)	AD	No intervention
				IG1	53 %	83.2 (8.3)	6.16 (6.16)		Conversation while walking (5x/week, 30min)
				IG2	60 %	81.7 (7.4)	5.44 (5.98)		Conversation (5x/week, 30min)
				All	53 %	82 (8)	6 (6), range: 0-21		16 weeks
Dawson et al. (2019)	RCT	n_r=23 n_a=23	Community-dwelling	CG	70 %	74.0 (10.4)	22.0 (3.1)	n.r.	No intervention
				IG	46 %	73.8 (8.5)	19.9 (6.1)		Moderate-intensity home-based functional exercise (strength, balance; 2x/week)
				All	57 %	73.9 (9.1)	20.8 (5.0), range: 9-28		12 weeks

Francese et al. (1997)	RCT	n_r=12 n_a=11	Care facility	CG	n.r.	n.r.	n.r.	AD	Watching a music video sing-a-along (3x/week, 20min)
				IG	n.r.	n.r.	n.r.		Exercise (3x/week, 20min)
				All	n.r.	n.r.	n.r.		7 weeks
Hauer et al. (2017)	RCT	n_r=34 n_a=28	Post-ward geriatric rehabili-tation	CG	71 %	83.3 (5.7)	18.2 (4.4)	n.r.	Usual care
				IG	65 %	81.4 (6.6)	19.5 (4.6)		Postural control and strength home training (at least once per day)
				All	n.r.	81.9 (5.7)	18.8 (4.7), range: n.r.		6 weeks
Hauer et al. (2012)	RCT	n_r=122 n_a=107	Outpa-tient geriatric rehabili-tation	CG	73 %	82.9 (7.0)	21.9 (3.2)	AD, VD, other	Motor placebo group training (2x/week, 60min)
				IG	74 %	82.3 (6.6)	21.7 (2.8)		Progressive resistance and func-tional training (2x/week, 120min)
				All	n.r.	n.r.	n.r., range: n.r.		3 months
Henskens et al. (2018)	RCT	n_r=87 n_a=87/ 16*	Care facility	CG1	77 %	84.7 (4.6)	10.2 (5.7)	AD, VD, MD, other/ unknown	Social activity
				CG2	77 %	85.1 (4.6)	12.1 (6.4)		Multicomponent aerobic and strength exercise (3x/week, 30-45min)
				IG1	91 %	86.1 (5.9)	13.2 (3.7)		Activities of daily living training + social activity
				IG2	63 %	87.0 (7.2)	13.6 (5.6)		Activities of daily living training + multicomponent aerobic and strength exercise (3x/week, 30-45min)
				All	n.r.	n.r.	n.r., range: n.r.		6 months
Kam-pragkou et al. (2017)	RCT	n_r=36 n_a=30	Hospital-ised	CG	n.r.	n.r.	16.0 (3.0)	AD	Memory games (3x/week, 40min)
				IG	n.r.	n.r.	14.7 (3.1)		Aerobic exercise (3x/week, 30min), memory games, attention, speech, and music (3x/week, 10min)
				All	n.r.	n.r.	n.r., range: n.r.		12 weeks

Kemoun et al. (2010)	RCT	n_r=38 n_a=31	Care facility	CG	73 %	81.7 (5.1)	12.9	AD	No intervention
				IG	75 %	82.0 (5.8)	12.6		Exercise (walking, equilibrium, stamina; 3x/week, 60min)
				All	74 %	81.8 (5.3)	n.r., range: 7-20		15 weeks
M.-J. Kim et al. (2016)	RCT	n_r=38 n_a=31	Care facility	CG	86 %	80.9 (6.1)	16.6 (4.0)	AD	Multicomponent intervention (5x/week, 2x/day, 60min)
				IG	68 %	81.9 (7.0)	13.4 (4.2)		Lower-limb aerobic exercise (5x/week, 60min) + multicomponent intervention (5x/week, 2x/day, 60min)
				All	76 %	81.5 (6.6)	14.8 (4.4), range: n.r.		6 months
Kovács et al. (2013)	RCT	n_r=86 n_a=62	Care facility	CG	79 %	79.3 (12.7)	20.9 (3.8)	n.r.	Usual care including social activities
				IG	83 %	76.4 (9.4)	20.9 (3.2)		Exercise (strengthening, balance, walking; 2x/week)
				All	81 %	n.r.	n.r., range: n.r.		12 months
Lam, Liao et al. (2018)	RCT	n_r=54 n_a=54	Day-care facility	CG	78 %	79.9 (6.7)	15.6 (4.5)	n.r.	Routine program including 30-60min of exercise (2x/week)
				IG	70 %	79.7 (5.5)	13.6 (4.7)		Routine program including 30-60min of exercise + whole-body vibration training (4-6min + rest, 2x/week)
				All	74 %	79.8 (6.1)	n.r., range: n.r.		9 weeks
Miu et al. (2008)	RCT	n_r=85 n_a=82	Community-dwelling	CG	63 %	78 (6)	Median=20	AD, VD, MD, PD	Conventional medical treatment
				IG	43 %	75 (7)	Median=20		Aerobic exercise (treadmill, bicycle, arm ergometer, flexibility exercises; 2x/week, 45-60min)
				All	54 %	76 (6)	Median=20, range: n.r.		12 weeks

Netz et al. (2007)	RCT	n_r=29 n_a=24	Day-care facility	CG	n.r.	n.r.	n.r.	n.r.	Social activity
				IG	n.r.	n.r.	n.r.		Group physical activity (2x/week, 45min)
				All	52 %	76.9 (6.7)	13.3 (5.8), range: n.r.		12 weeks
Padala et al. (2017)	RCT	n_r=30 n_a=30	Commu-nity-dwelling	CG	40 %	73.9 (7.1)	22.7 (2.3)	AD	Walking (5x/week, 30min)
				IG	33 %	72.1 (5.3)	23.3 (2.2)		Wii-Fit (5x/week, 30min)
				All	37 %	73.0 (6.2)	22.9 (2.2), range: n.r.		8 weeks
Padala et al. (2012)	RCT	n_r=22 n_a=22	Care facility	CG	73 %	81.6 (5.2)	24.9 (3.6)	AD	Walking (5x/week, 30min)
				IG	73 %	79.3 (9.8)	22.6 (4.3)		Wii-Fit (5x/week, 30min)
				All	73 %	n.r.	n.r., range: n.r.		8 weeks
Pedrinolla et al. (2018)	RCT	n_r=53 n_a=34	Commu-nity-dwelling	CG	61 %	79 (6)	21 (5)	AD	Standard cognitive treatment (3x/week, 90min)
				IG	63 %	80 (7)	22 (5)		Aerobic and strength training (3x/week, 90min)
				All	62 %	n.r.	n.r., range: n.r.		6 months
Pitkälä, Pöysti et al. (2013)	RCT	n_r=210 n_a=194	Commu-nity-dwelling	CG	37 %	78.1 (5.3)	17.7 (6.2)	AD	Usual community care
				IG1	36 %	78.3 (5.1)	18.5 (6.3)		Group-based exercise (2x/week, 60min)
				IG2	43 %	77.7 (5.4)	17.8 (6.6)		Home-based exercise (2x/week, 60min)
				All	39 %	n.r.	n.r., range: n.r.		12 months
Pomeroy et al. (1999)	RCT	n_r=81 n_a=78	Commu-nity-dwelling	CG	74 %	81.8 (8.4)	n.r.	AD, MD, MID, undefined	Non-physical activities
				IG	74 %	82 (8.0)	n.r.		Physiotherapy (up to 10x 30min)
				All	74 %	81.9 (n.r.)	n.r., range: n.r.		2 weeks

Roach et al. (2011)	RCT	n_r=105 n_a=82	Care facility	CG	n.r.	88.2 (5.8)	9.4 (7.2)	AD	Conversation (5x/week, 15-30min)
				IG1	n.r.	89.2 (6.5)	8.7 (7.8)		Exercise (strength, flexibility, balance, walking; 5x/week, 25-50min)
				IG2	n.r.	87.3 (6.1)	12.2 (7.5)		Walking (5x/week, 15-30min)
				All	n.r.	88.2 (6.1)	10.2 (7.6), range: n.r.		16 weeks
Rolland et al. (2007)	RCT	n_r=134 n_a=110	Care facility	CG	79 %	83.1 (7.0)	7.9 (6.4)	AD	Routine medical care
				IG	72 %	82.8 (7.8)	9.7 (6.8)		Exercise (walking, strength, balance, flexibility; 2x/week, 60min)
				All	75 %	83 (7.4)	8.8 (6.6), range: n.r.		12 months
Santana-Sosa et al. (2008)	RCT	n_r=16 n_a=16	Care facility	CG	63 %	73 (4)	19.9 (1.7)	AD	Routine nursing/medical care
				IG	63 %	76 (4)	20.1 (2.3)		Training program (resistance, flexibility, joint mobility, balance/ coordination; 3x/week, 75min)
				All	63 %	n.r.	n.r., range: n.r.		12 weeks
Schwenk, Dutzi et al. (2014)	Quasi RCT	n_r=148 n_a=130	Geriatric hospital	CG	76 %	83.9 (6.1)	22.2 (2.3)	n.r.	Usual care treatment (occupational therapy, speech therapy, physiotherapy including exercise)
				IG	84 %	84.2 (6.2)	21.4 (2.6)		Usual care treatment (see above) + intensive exercise (2x daily, up to 60min)
				All	n.r.	n.r.	n.r., range: n.r.		During rehabilitation period, average: 18.1 (6.8) days
Schwenk, Zieschang et al. (2014)	RCT	n_r=61 n_a=49	Outpatients	CG	63 %	82.3 (7.9)	21.7 (2.9)	n.r.	Low-intensity motor placebo activity program (60min, 2x/week)
				IG	65 %	80.4 (7.1)	21.0 (2.9)		Progressive resistance and functional group training (120min, 2x/week)
				All	64 %	81.9 (7.5)	21.4 (2.9), range: n.r.		3 months

Schwenk et al. (2010)	RCT	n_r=61 n_a=49	n.r.	CG	63 %	82.3 (7.9)	21.7 (2.9)	n.r.	Unspecific low-intensity exercise (60min, 2x/week)
				IG	65 %	80.4 (7.1)	21.0 (2.9)		Dual-task–based exercise training (120min, 2x/week)
				All	64 %	81.9 (7.5)	21.4 (2.9), range: n.r.		12 weeks
Sobol et al. (2016)	RCT	n_r=200 n_a=189	Community-dwelling	CG	39 %	71.3 (7.3)	24.1 (3.8)	AD	Usual care
				IG	48 %	69.8 (7.4)	23.8 (3.4)		Moderate-to-high intensity aerobic exercise (3x/week, 60min)
				All	44 %	70.5 (7.4)	24.0 (3.6), range: n.r.		16 weeks
Souto Barreto et al. (2017)	Cluster RCT	n_r=98 n_a=91	Care facility	CG	77 %	86.9 (5.8)	10.8 (5.5)	AD, VD, MD	Social activity (2x/week, 60min)
				IG	93 %	88.3 (5.1)	11.4 (6.2)		Multicomponent training (range of motion, coordination, balance strength, aerobic; 2x/week, 60min)
				All	86 %	n.r.	n.r., range: n.r.		24 weeks
Steinberg et al. (2009)	RCT	n_r=27 n_a=27	Community-dwelling	CG	69 %	74.0 (8.1)	15.5 (5.4)	AD	Home safety assessment
				IG	71 %	76.5 (3.9)	20.1 (5.1)		Home-based exercise (aerobic, strength, balance, flexibility; daily)
				All	79 %	n.r.	n.r., range: n.r.		12 weeks
Suttanon et al. (2013)	RCT	n_r=40 n_a=40	Community-dwelling	CG	57 %	80.5 (6.0)	21.7 (4.4)	AD	Home-based education program
				IG	68 %	83.4 (5.1)	20.9 (4.7)		Home-based individually tailored exercise (balance, strength, walking; 5x/week)
				All	63 %	81.9 (5.7)	n.r., range: n.r.		6 months

Tappen et al. (2000)	RCT	n_r=71 n_a=65	Care facility	CG	n.r.	89.6 (6.5)	12.5 (5.9)	AD	Conversation (3x/week, 30min)
				IG1	n.r.	84.3 (7.5)	10.8 (6.0)		Walking combined with conversation (3x/week, 30min)
				IG2	n.r.	87.4 (5.9)	9.8 (6.0)		Walking (3x/week, 30min)
				All	84 %	87 (n.r.)	10.8 (n.r.), range: 0-23		16 weeks
Telenius et al. (2015a)	RCT	n_r=170 n_a=160	Care facility	CG	75 %	86.4 (7.8)	15.8 (5.0)	n.r.	Leisure activities (2x/week, 50-60min)
				IG	72 %	86.9 (7)	15.6 (5.0)		Intensive strengthening and balance exercises (2x/week, 50-60min)
				All	74 %	86.7 (7.4)	15.7 (5.0), range: n.r.		12 weeks
Toots et al. (2017)	Cluster RCT	n_r=186 n_a=153	Care facility	CG	76 %	85.9 (7.8)	14.4 (3.5)	AD, VD, MD, other	Seated attention control activity
				IG	75 %	84.4 (6.2)	15.4 (3.4)		High-intensity functional exercise (strength, balance, mobility; 2-3x/week, 45min)
				All	76 %	85.1 (7.1)	14.9 (3.5), range: n.r.		4 months
Toots et al. (2016)	Cluster RCT	n_r=186 n_a=167	Care facility	CG	76 %	85.9 (7.8)	14.4 (3.5)	AD, VD, MD, other	Seated attention control activity
				IG	75 %	84.4 (6.2)	15.4 (3.4)		High-intensity functional exercise (strength, balance, mobility; 2-3x/week, 45min)
				All	76 %	85.1 (7.1)	14.9 (3.5), range: n.r.		4 months
Toulotte et al. (2003)	RCT	n_r=20 n_a=20	n.r.	CG	n.r.	81.9 (4.1)	18.0 (5.4)	AD, PD, stroke, unknown	Daily routine
				IG	n.r.	81.0 (5.6)	14.7 (7.6)		Physical training (2x/week, 45min)
				All	n.r.	81.4 (4.7)	16.3 (6.5), range: n.r.		16 weeks

Venturelli et al. (2011)	RCT	n_r=24 n_a=21	Care facility	CG	n.r.	85 (5)	12 (2)	AD	Routine care
				IG	n.r.	83 (6)	13 (2)		Walking program (4x/week, 30min)
				All	n.r.	84 (5)	n.r., range: n.r.		24 weeks
Vreugdenhil et al. (2012)	RCT	n_r=40 n_a=40	Community-dwelling	CG	75 %	74.7 (n.r.)	21.0 (6.3)	AD	Usual treatment
				IG	45 %	73.5 (n.r.)	22.9 (5.0)		Community-based home exercise program (daily)
				All	60 %	74.1 (n.r.)	22.0 (n.r.), range 10-28		4 months
Werner et al. (2017)	RCT	n_r=97 n_a=80	Geriatric hospital, care facility, community-dwelling	CG	78 %	82.5 (5.7)	21.5 (3.0)	n.r.	Low-intensity motor placebo activity (2x/week, 60min)
				IG	71 %	82.6 (6.1)	22.2 (2.9)		Motor learning exercise program on compensatory sit-to-stand maneuvers (2x/week, 90min)
				All	74 %	82.5 (5.9)	21.9 (2.9), range: n.r.		10weeks
Wesson et al. (2013)	RCT	n_r=22 n_a=21	Community-dwelling	CG	36 %	80.9 (5.0)	22.5 (4.3)	n.r.	Usual care
				IG	46 %	78.7 (4.2)	24.5 (3.1)		Strength and balance training exercises and home hazard reduction
				All	41 %	n.r.	n.r., range: 15-29		12 weeks
Wiloth et al. (2018)	RCT	n_r=99 n_a=84	Geriatric hospital, care facility, community-dwelling	CG	74 %	82.2 (5.3)	21.7 (2.9)	n.r.	Supervised motor placebo group training (2x/week, 60min)
				IG	70 %	82.7 (6.2)	22.2 (2.8)		Comprehensive motor-cognitive intervention program (game-based training using Physiomat, dual-task training, motor learning exercise program) (2x/week, 90min)
				All	72 %	82.9 (5.8)	22.0 (2.9), range: n.r.		10 weeks

Yoon et al. (2013)	RCT	n_r=30 n_a=20	Care facility	CG	n.r.	70.1 (12.2)	18.7 (1.2)	n.r.	Conventional physical therapy (5x/week, 30min) + cognitive activity (3x/week, 20min)
				IG	n.r.	77.9 (7.5)	18.0 (1.5)		Conventional physical therapy (5x/week, 30min) + cognitive activity combined with cycling (3x/week, 20min)
				All	n.r.	n.r.	n.r., range: 16-23		12 weeks

AD: Alzheimer's disease, CG: control group, CI_{95}: 95 % confidence interval, FLD: Frontal lobe dementia, IG: intervention group, LBD: Lewy body disease, MD: mixed dementia (AD+VD), MID: multi-infarct dementia, MMSE: Mini-Mental State Examination, n_a: number of analysed participants, n_r: number of randomised participants, n.r.: not reported, PD: Parkinson's disease dementia, RCT: randomised controlled trial, SD: standard deviation, VD: vascular dementia;

* Intention-to-treat analysis and complete-case analysis

Additional file 6

Study characteristics main search

***Table B.** Study characteristics main search

Reference	**Sample size**	**Setting**	**Sample characteristics**				**Cueing**	**Psychometric property (time interval if applicable)**	**Methodology**	**COSMIN risk of bias**	
			Gender [% female]	**Age [years] Mean (SD)**	**MMSE Mean (SD), range**	**Aetiology of dementia**				**Relative reliability**	**Absolute reliability**
Alencar et al. (2012)	n=76	Non-institutionalised	84 %	83.9 (5.8)	12.7 (7.2), n.r. SG0: n.r. SG1: n.r. SG2: n.r. SG3: n.r.	AD, VD	(a) n.r.	Between-day TRR (1 week)	ICC (n.r.)	Adequate	N/A

Blankevoort et al. (2013)	n=58	(Day-)care facilities	71 %	82.5 (5.3)	19.2 (4.4), 10-28 SG1: 22.8 (2.1), 20-28 SG2: 15.5 (2.6), 10-19	n.r.	(d) verbal & visual/tactile cueing	Between-day TRR (1 week)	ICC (2-way, random, absolute agreement on single measures model) with CI_{95} SEM with CI_{95}, MDC_{95}	Adequate	Adequate
Bossers, van der Woude et al. (2014)	n=42	Care facilities	79 %	86.7 (5.2)	17.1 (4.3), 9-24 SG1: n.r., 20-24 SG2: n.r., 9-19	AD, VD, MD, LBD	(b) no cueing	Between-day TRR (1 week)	ICC (2-way mixed single measure, absolute agreement model) with CI_{95} SEM with CI_{95}, MDC_{95}	Adequate	Adequate
Bronas et al. (2017)	n=44	Community-dwelling	45 %	78.4 (6.8)	21.6 (3.3), n.r.	AD	(a) n.r./(c) verbal cueing	Simple associations	Pearson bivariate correlation	N/A	N/A
Graessel et al. (2009)	n=46	Care facility	91 %	85.9 (6.6)	16.0 (6.1), 1-28	n.r.	(b) no cueing	Construct validity	Spearman's rank correlation coefficient	N/A	N/A
								Internal consistency	Cronbach's α, Spearman's rank correlation coefficient	N/A	N/A
								Between-day TRR (2 weeks)	Spearman's rank correlation coefficient	Doubtful	N/A

H.-S. Lee et al. (2017)	n=53	Care facility	n.r.	83.8 (9.9)	13.8 (5.7), n.r.	n.r.	(e) more extensive cueing than (c) & (d) including physical assistance	IRR (3 raters)	ICC (2,3) SEM, MDC_{95}, $MDC_{\%}$	Adequate	Very good
								Between-day TRR (3-7 days)		Adequate	Adequate
Luttenberger et al. (2012)	n=139	Care facility	83 %	84.7 (4.9)	15.2 (5.3), n.r.	n.r.	(b) no cueing	Construct validity	Spearman correlation coefficient	N/A	N/A
								Criterion-related validity	Correlations Kruskal-Wallis test Mann-Whitney U-test	N/A	N/A
								Internal consistency	Cronbach's alpha, Spearman correlation coefficient	N/A	N/A
McGough et al. (2013)	n=31 (8)*	Care facility	94 %	83.6 (7.0)	12.4 (7.0), 2-26	n.r.	(c) verbal cueing	Concurrent validity	Bivariate and partial correlations	N/A	N/A
								Within-day TRR (4 h)	ICC (2-way, random)	Adequate	N/A
Muir-Hunter et al. (2015)	n=15	Community-dwelling (day program)	27 %	80.2 (5.0)	20.0 (5.5), n.r.	AD	(e) more extensive cueing than (c) & (d) including physical assistance	IRR	ICC (n.r.) with CI_{95}	Adequate	Very good
								Between-day TRR (1 week)	SEM, MDC_{95}	Adequate	Adequate

Ries et al. (2009)	n=51	(Day-)care facility	67 %	80.7 (8.8)	13.1 (8.2), n.r. SG1/2: 17.4 (4.5), 10-26 SG3: 10.2 (8.8), n.r.	AD	(e) more extensive cueing than (c) & (d) including physical assistance	Within-day TRR (30-60 min)	ICC (2,2)/ICC (2,1) SEM, MDC_{90}	Adequate	Adequate
Schwenk, Hauer et al. (2014)	n=77	Geriatric hospital	70 %	81.8 (6.3)	22.1 (3.2), n.r.	n.r.	(a) n.r.	Criterion validity	Mann-Whitney U-test Logistic regression analysis	N/A	N/A
Suttanon et al. (2011)	n=14	Community-dwelling	50 %	79.6 (6.2)	21.4 (5.0), n.r.	AD	(d) verbal & visual/tactile cueing	Between-day TRR (1 week)	ICC (3,1) SEM, MDC_{95}, CV	Adequate	Adequate
Suzuki et al. (2009)	n=60	Care facility	77 %	86.6 (6.2)	n.r., 0-23 SG1/2: n.r., 11-23 SG3: n.r., 0-10	AD	(a) n.r.	Known group validity	Mann-Whitney U-test/unpaired t-test	N/A	N/A
								Predictive validity	Logistic regression analysis	N/A	N/A
								Within-session TRR (3min)	ICC (n.r.)	Adequate	N/A
Suzuki et al. (2012)	n=54	Care facility	76 %	87.0 (5.7)	n.r., 0-24	AD, MD	(a) n.r.	Predictive validity	Chi^2 test/Mann-Whitney U-test/ unpaired t-test Forward stepwise logistic regression analysis	N/A	N/A

Tappen et al. (1997)	n=33	Care facility	64 %	84.7 (3.9)	9.3 (6.0), 0-24	AD	(e) more extensive cueing than (c) & (d) including physical assistance	IRR	ICC (2,1)	Adequate	N/A
								Between-day and within-day TRR (1 week, AM/PM)	ICC (3,1)	Adequate	N/A
Telenius et al. (2015b)	n=33	Care facility	76 %	82.7 (7.2)	15.8 (5.4), n.r.	n.r.	(e) more extensive cueing than (c) & (d) including physical assistance	Internal consistency	Cronbach's α, item-to-total correlation	N/A	N/A
								IRR	Weighted κ ICC model 2.1 SEM, MDC_{95}, $MDC_{\%}$	Adequate	Very good
Thomas and Hageman (1999)	n=12	Day-care facility	100 %	80.5 (6.2)	16.9 (7.3), n.r.	Various	(d) verbal & visual/tactile cueing	Between-day TRR (7 days)	ICC (n.r.)	Adequate	N/A
Werner et al. (2018)	n=97	Geriatric hospital, care facility, community-dwelling	74 %	82.5 (5.9)	21.0 (2.9), n.r.	n.r.	(b) no cueing	Concurrent validity	Point-biserial correlation coefficients	N/A	N/A
								Intra-rater reliability (repeated baseline scoring 4 weeks later)	Percentage agreement, Cohen's κ ICC (3,1) with CI_{95}	Very good	N/A
								IRR	Percentage agreement, Cohen's κ ICC (2,1) with CI_{95}	Very good	N/A

Wiloth et al. (2016)	n=105 (74)*	Geriatric hospital, care facility, community-dwelling	72 %	82.7 (5.9)	21.9 (2.8), n.r.	n.r.	(b) no cueing	Construct validity	Spearman's rank correlations	N/A	N/A
								Between-day TRR (2-5 days)	Spearman's rank correlations ICC (2-way mixed model) with CI_{95}	Adequate	N/A
Wittwer et al. (2008)	n=20	Community-dwelling	50 %	80.6 (5.2)	22.0 (3.5), 13-27	AD	(b) no cueing	Between-day TRR (1 week)	ICC (3,1) with CI_{95} MDC_{95}, CV	Adequate	Adequate
Wittwer et al. (2013)	n=16	Community-dwelling	63 %	81.1 (5.2)	21.4 (4.0), 13-26	AD	(a) n.r.	Between-day TRR (1 week)	ICC (3,1) with CI_{95} SEM, MDC_{95}	Adequate	Adequate

AD: Alzheimer's disease, AM/PM: morning/afternoon measures, CI_{95}: 95 % confidence interval, COSMIN: COnsensus-based Standards for the selection of health Measurement INstruments, CV: coefficient of variation, ICC: intraclass correlation coefficient, IRR: inter-rater reliability, LBD: Lewy body disease, MD: mixed dementia (AD+VD), MDC: minimal detectable change with MDC_{95}=SEM*1.96*Sqrt(2), MDC_{90}=SEM*1.65*Sqrt(2), $MDC_{\%}$=(MDC_{95} / mean)*100, MMSE: Mini-Mental State Examination, n: number of participants, N/A: not applicable, n.r.: not reported, SD: standard deviation, SEM: standard error of measurement with SEM=SD*Sqrt(1-ICC), SEM=SD*Sqrt(1-r), SG0: subgroup borderline, SG1: subgroup mild dementia, SG1/2: subgroup mild to moderate dementia, SG2: subgroup moderate dementia, SG3: subgroup severe dementia, TRR: test-retest reliability, VD: vascular dementia

* Reliability was assessed in a subgroup

Additional file 7

Data extraction first search

***Table C.** Data extraction first search

Reference	Motor assessments		CG			IG			Statistics / p-value	Effect size (reported)	Time*group interaction effect size (calculate)
			Baseline Mean (SD)	Post Mean (SD)	Difference Mean (SD)	Baseline Mean (SD)	Post Mean (SD)	Difference Mean (SD)			
Aguiar et al. (2014)	TUG [s]		14.5 (4.9)	12.9 (4.2)	-1.6 (2.4)	15.0 (5.2)	15.0 (5.6)	0.0 (1.7)	p=0.062 [a]	n.r.	0.79
Arcoverde et al. (2014)	FR [cm]		19.0 (4.2)	18.3 (n.r.)	-0.7 (4.3)	20.0 (4.3)	25.1 (n.r.)	5.1 (2.9)	p=0.00 [a]	1.48 [y]	1.67
	BBS		n.r. (n.r.)	n.r. (n.r.)	n.r.	n.r. (n.r.)	n.r. (n.r.)	n.r.	p=0.00 [b]	1.04 [y]	N/A
	TUG [m/s]		n.r. (n.r.)	n.r. (n.r.)	n.r. (n.r.)	n.r. (n.r.)	n.r. (n.r.)	n.r. (n.r.)	p=0.00 [b]	1.58 [y]	N/A
	Cognitive TUG [m/s]		n.r. (n.r.)	n.r. (n.r.)	n.r. (n.r.)	n.r. (n.r.)	n.r. (n.r.)	n.r. (n.r.)	p=0.24 [a]	1.03 [y]	N/A
	30s CST		9.0 (3.0)	8.5 (n.r.)	-0.5 (0.9)	9.0 (2.0)	10.0 (n.r.)	1.0 (2.4)	p=0.08 [a]	0.50 [y]	0.87
Bossers et al. (2015)	FICSIT-4		2.3 (1.4)	2.0 (1.4)	n.r.	IG1: 2.3 (1.0) IG2: 2.8 (1.2)	IG1: 2.8 (0.9) IG2: 2.5 (1.1)	n.r.	F(2,105)=5.36, p=0.024 [c]	CG vs. IG1: 0.30 [y] CG vs. IG2: 0.08 [y] IG1 vs. IG2: 0.33 [y]	N/A
	Figure of Eight Test	walking speed [m/s]	0.4 (0.3)	0.3 (0.3)	n.r.	IG1: 0.3 (0.2) IG2: 0.4 (0.3)	IG1: 0.4 (0.3) IG2: 0.4 (0.4)	n.r.			N/A
		oversteps	9.0 (8.0)	9.7 (8.2)	n.r.	IG1: 7.7 (7.6) IG2: 5.9 (7.4)	IG1: 7.3 (7.5) IG2: 7.9 (7.7)	n.r.			N/A

GMWT	time [s]	21.5 (12.7)	21.2 (13.7)	n.r.	IG1: 23.3 (12.7) IG2: 18.7 (7.7)	IG1: 21.6 (11.4) IG2: 19.0 (8.9)	n.r.			N/A
	oversteps	2.6 (2.9)	2.7 (2.7)	n.r.	IG1: 2.1 (2.1) IG2: 0.8 (1.6)	IG1: 1.7 (2.4) IG2: 1.1 (1.6)	n.r.			N/A
TUG [s]		27.6 (18.8)	27.7 (19.2)	n.r.	IG1: 23.0 (13.0) IG2: 24.3 (14.0)	IG1: 20.4 (9.2) IG2: 23.8 (15.0)	n.r.	F(2,105)=1.28, p=0.282 [c]	CG vs. IG1: 0.28 [y] CG vs. IG2: 0.06 [y] IG1 vs. IG2: 0.26 [y]	N/A
6m WT	walking speed [m/s]	0.7 (0.3)	0.6 (0.3)	n.r.	IG1: 0.7 (0.3) IG2: 0.8 (0.3)	IG1: 0.8 (0.3) IG2: 0.7 (0.4)	n.r.			N/A
	step length [m]	0.4 (0.1)	0.4 (0.1)	n.r.	IG1: 0.4 (0.1) IG2: 0.4 (0.1)	IG1: 0.4 (0.1) IG2: 0.4 (0.1)	n.r.			N/A
Modified 30s CST		6.2 (4.8)	5.4 (6.3)	n.r.	IG1: 6.8 (3.4) IG2: 7.1 (4.4)	IG1: 8.2 (3.6) IG2: 6.3 (4.8)	n.r.	F(2,105)=7.07, p=0.004 [c]	CG vs. IG1: 0.38 [y] CG vs. IG2: 0.04 [y] IG1 vs. IG2: 0.36 [y]	N/A
Maximum knee extension strength (dynamometer) [N]		218.9 (84.1)	203.2 (74.0)	n.r.	IG1: 205.9 (91.0) IG2: 196.0 (99.0)	IG1: 208.8 (85.6) IG2: 186.1 (84.3)	n.r.			N/A

	6min WT [m]		229.5 (136.4)	221.8 (159.5)	n.r.	IG1: 217.6 (90.3) IG2: 231.5 (136.4)	IG1: 267.2 (101.2) IG2: 235.5 (148.7)	n.r.	F(2,105)=4.53, p<0.049 [c]	CG vs. IG1: 0.47 [y] CG vs. IG2: 0.08 [y] IG1 vs. IG2: 0.38 [y]	N/A
Bossers et al. (2016)	Physical Performance Test		12.9 (5.9)	11.2 (6.4)	n.r.	IG1: 12.9 (3.4) IG2: 13.4 (5.1)	IG1: 14.2 (3.5) IG2: 13.2 (4.9)	n.r.	$Chi^2(2)=11.93$, p=0.003 [d]	CG vs. IG1: 0.62 [y] CG vs. IG2: 0.29 [y] IG1 vs. IG2: 0.36 [y]	N/A
	E-ADL Test		26.9 (3.4)	25.3 (5.1)	n.r.	IG1: 26.5 (3.7) IG2: 26.8 (3.5)	IG1: 28.1 (2.7) IG2: 27.3 (3.8)	n.r.	$Chi^2(2)=16.40$, p<0.001 [d]	CG vs. IG1: 0.85 [y] CG vs. IG2: 0.53 [y] IG1 vs. IG2: 0.31 [y]	N/A
Burgener et al. (2008)	Single leg stance [s]	left leg, eyes closed	1.7 (1.1)	3.6 (4.8)	1.9 (n.r.)	3.4 (6.8)	6.3 (14.2)	2.9 (n.r.)	IG: n.s. [e] p=0.62 [f]	n.r.	N/A
		right leg, eyes open	6.0 (5.5)	3.7 (2.1)	-2.3 (n.r.)	5.9 (5.6)	10.4 (15.5)	4.5 (n.r.)	IG: n.s. [e] p=0.09 [f]	n.r.	N/A
	BBS		50.8 (4.2)	50.5 (3.5)	-0.3 (n.r.)	49.1 (5.0)	50.8 (4.3)	1.7 (n.r.)	p=0.87 [f]	n.r.	N/A

Cancela et al. (2016)	TUG [s]		23.4 (6.9)	n.r.	-0.6 (n.r.) -1.8 (n.r.) *	24.0 (10.1)	n.r.	-2.1 (n.r.) -3.0 (n.r.) *	F(1,187)=5.43, p=0.03 [g] F(1,111)=4.10, p=0.04 [h] *	n.r.	0.35 0.38 *
Christo-foletti et al. (2008)	BBS		35.2 (2.6)	27.4 (3.2)	n.r.	IG1: 39.5 (1.9) IG2: 37.4 (2.0)	IG1: 41.7 (2.4) IG2: 37.7 (2.8)	n.r.	CG vs. IG1: F=10.3, p<0.05 [i] CG vs. IG2: F=7.9, p<0.05 [i]	n.r.	N/A
	TUG	time [s]	30.6 (6.5)	35.6 (8.6)	n.r.	IG1: 13.7 (1.2) IG2: 22.3 (4.4)	IG1: 12.9 (1.0) IG2: 22.1 (4.0)	n.r.	n.s. [i]	n.r.	N/A
		steps	31.3 (4.2)	35.3 (6.4)	n.r.	IG1: 19.9 (1.4) IG2: 28.2 (3.6)	IG1: 18.3 (1.2) IG2: 25.5 (3.6)	n.r.	n.s. [i]	n.r.	N/A
Cott et al. (2002)	2-min walk test [m]		48.0 (28.8)	47.7 (33.8)	n.r.	IG1: 52.8 (27.6) IG2: 52.6 (24.2)	IG1: 53.3 (27.5) IG2: 56.4 (34.4)	n.r.	n.s. [e, j]	n.r.	N/A
Dawson et al. (2019)	Modified BBS		38.5 (8.0)	36.6 (8.7)	n.r.	39.5 (3.3)	41.5 (2.2)	n.r.	B=4.0, β=0.3, t=4.1, p=0.001 [kl]	n.r.	N/A
	8-foot walk test [m/s]	comforta-ble pace	0.7 (0.2)	0.6 (0.3)	n.r.	0.7 (0.2)	0.7 (0.1)	n.r.	B=0.01, β=0.2, t=0.6, p=0.6 [k]	n.r.	N/A
		fast pace	1.4 (0.6)	1.3 (0.6)	n.r.	1.2 (0.3)	1.6 (0.3)	n.r.	B=0.3, β=0.4, t=2.6, p=0.02 [k]	n.r.	N/A
	30s CST		15.7 (6.1)	13.2 (4.9)	n.r.	14.0 (5.8)	17.9 (6.8)	n.r.	B=5.9, β=0.5, t=3.3, p=0.004 [k]	n.r.	N/A

Francese et al. (1997)	POMA		1.8 (2.1)	0.4 (0.9)	n.r.	3.0 (2.8)	8.7 (4.3)	n.r.	t(10)=2.00, p<.05 [f] CG: t(4)=-1.00, p≥.05 [e] IG: t(5)=3.00, p=0.05 [e]	n.r.	N/A
	Physical therapy assessment		38.8 (34.7)	43.6 (37.7)	n.r.	63.8 (18.3)	89.7 (10.0)	n.r.	t(10)=3.20, p=0.01 [f] CG: t(4)=0.83, p≥0.05 [e] IG: t(5)=4.33, p=0.01 [e]	n.r.	N/A
Hauer et al. (2017)	POMA	total score	18.4 (7.0)	18.6 (5.2)	n.r.	18.6 (7.5)	23.0 (4.4)	n.r.	p=0.006 [i]	0.25 [z]	N/A
		balance score	11.5 (3.4)	11.7 (1.8)	n.r.	11.0 (3.4)	13.3 (1.7)	n.r.	p=0.034 [i]	0.16 [z]	N/A
		gait score	6.9 (3.9)	6.9 (3.9)	n.r.	7.6 (4.5)	9.7 (2.8)	n.r.	p=0.019 [i]	0.19 [z]	N/A
	Body-fixed-sensor-based STS analysis (DynaPort): sit-to-stand performance	duration [s]	2.0 (0.6)	2.5 (0.9)	n.r.	1.8 (0.4)	1.4 (0.2)	n.r.	p=0.064 [i]	0.30 [z]	N/A
		hip flexion, duration [s]	1.0 (0.4)	1.1 (0.4)	n.r.	0.9 (0.2)	0.8 (0.2)	n.r.	p=0.451 [i]	0.06 [z]	N/A
		hip extension, duration [s]	0.9 (0.2)	1.4 (0.6)	n.r.	0.9 (0.3)	0.6 (0.1)	n.r.	p=0.018 [i]	0.44 [z]	N/A
		hip flexion, max. angular velocity [°/s]	74.0 (21.1)	60.4 (10.0)	n.r.	86.2 (35.9)	106.0 (54.5)	n.r.	p=0.239 [i]	0.14 [z]	N/A

	hip extension, max. angular velocity [°/s]	46.9 (26.5)	41.4 (15.4)	n.r.	42.2 (22.7)	37.7 (17.5)	n.r.	p=0.919 [i]	0.001 [z]	N/A
Body-fixed-sensor-based STS analysis (DynaPort): stand-to-sit performance	duration [s]	2.0 (0.8)	2.4 (0.7)	n.r.	2.2 (0.5)	1.5 (0.3)	n.r.	p=0.014 [i]	0.47 [z]	N/A
	hip flexion, duration [s]	1.0 (0.4)	1.3 (0.3)	n.r.	1.1 (0.3)	0.7 (0.2)	n.r.	p=0.015 [i]	0.46 [z]	N/A
	hip extension, duration [s]	0.9 (0.4)	1.1 (0.4)	n.r.	1.0 (0.2)	0.8 (0.2)	n.r.	p=0.044 [i]	0.35 [z]	N/A
	hip flexion, max. angular velocity [°/s]	36.8 (17.7)	31.5 (12.4)	n.r.	41.1 (13.8)	46.5 (27.1)	n.r.	p=0.369 [i]	0.08 [z]	N/A
	hip extension, max. angular velocity [°/s]	81.8 (39.0)	52.5 (10.3)	n.r.	75.3 (22.6)	107.7 (50.1)	n.r.	p=0.006 [i]	0.55 [z]	N/A
SPPB	total score	4.4 (2.9)	4.7 (2.4)	n.r.	5.0 (2.7)	7.0 (2.7)	n.r.	p=0.010 [i]	0.23 [z]	N/A
	chair rise score	0.7 (1.3)	0.7 (1.1)	n.r.	0.6 (0.8)	1.5 (1.4)	n.r.	p=0.007 [i]	0.25 [z]	N/A
	balance score	2.6 (.7)	2.7 (.7)	n.r.	2.9 (.9)	3.6 (0.7)	n.r.	p=0.066 [i]	0.12 [z]	N/A
	gait score	1.1 (1.4)	1.3 (1.1)	n.r.	1.4 (1.3)	1.9 (1.3)	n.r.	p=0.395 [i]	0.03 [z]	N/A
	8-foot walk test [m/s]	0.4 (0.2)	0.4 (0.2)	n.r.	0.5 (0.2)	0.6 (0.2)	n.r.	p=0.153 [i]	0.12 [z]	N/A
	5x STS [s]	19.1 (9.8)	24.4 (12.0)	n.r.	20.9 (5.7)	16.0 (4.0)	n.r.	p=0.009 [i]	0.45 [z]	N/A

Hauer et al. (2012)	POMA	total score	19.8 (5.4)	20.6 (6.0)	n.r.	20.1 (4.8)	24.5 (3.7)	n.r.	p<0.001 [c]	0.22 [aa]	N/A
		balance score	10.8 (3.1)	11.1 (3.2)	n.r.	10.8 (2.8)	13.2 (1.9)	n.r.	p<0.001 [c]	0.23 [aa]	N/A
		gait score	9.0 (2.7)	9.4 (3.1)	n.r.	9.3 (2.3)	11.4 (1.9)	n.r.	p<0.001 [c]	0.19 [aa]	N/A
	TUG [s]		17.9 (16.0)	17.5 (17.3)	n.r.	14.9 (6.7)	11.2 (4.5)	n.r.	p=0.009 [c]	0.06 [aa]	N/A
	Handgrip dynamometer [KPa]		59.7 (16.6)	59.7 (15.7)	n.r.	59.1 (17.8)	60.9 (17.4)	n.r.	p=0.55 [c]	0.004 [aa]	N/A
	One-repetition maximum in leg press [kg]		140.9 (44.0)	136.5 (45.4)	n.r.	148.7 (57.9)	225.2 (79.7)	n.r.	p<0.001 [c]	0.43 [aa]	N/A
	Maximum isometric strength (dynamom-eter)	knee ex-tension [N]	65.8 (24.8)	66.8 (25.5)	n.r.	68.3 (27.3)	81.3 (27.4)	n.r.	p<0.001 [c]	0.12 [aa]	N/A
		knee ex-tension [Ns]	267.7 (101.9)	270.5 (103.0)	n.r.	277.2 (114.4)	324.6 (122.6)	n.r.	p=0.001 [c]	0.09 [aa]	N/A
		knee flex-ion [N]	31.4 (12.1)	34.9 (12.9)	n.r.	33.3 (12.6)	43.7 (14.5)	n.r.	p<0.001 [c]	0.15 [aa]	N/A
		knee flex-ion [Ns]	137.2 (52.7)	143.5 (54.9)	n.r.	145.1 (55.3)	178.2 (61.3)	n.r.	p<0.001 [c]	0.14 [aa]	N/A
		ankle flex-ion [N]	52.3 (21.8)	52.5 (23.7)	n.r.	56.5 (24.7)	63.1 (26.1)	n.r.	p=0.01 [c]	0.06 [aa]	N/A
		ankle flex-ion [Ns]	212.6 (93.2)	214.3 (99.0)	n.r.	229.0 (100.3)	257.1 (107.5)	n.r.	p=0.01 [c]	0.06 [aa]	N/A
	Gait perfor-mance	walking sped [m/s]	0.9 (0.3)	1.0 (0.3)	n.r.	0.9 (0.3)	1.2 (0.4)	n.r.	p<0.001 [c]	0.28 [aa]	N/A
		step length [m]	0.5 (0.1)	0.5 (0.1)	n.r.	0.5 (0.2)	0.6 (0.1)	n.r.	p<0.001 [c]	0.16 [aa]	N/A
		cadence [steps/min]	116.7 (18.9)	117.9 (20.7)	n.r.	117.1 (18.7)	131.7 (17.1)	n.r.	p<0.001 [c]	0.18 [aa]	N/A

	5x STS [s]		17.6 (9.3)	19.7 (15.9)	n.r.	17.3 (6.8)	11.8 (3.2)	n.r.	p<0.001 [c]	0.15 [aa]	N/A
	Stair-climbing performance [s]		16.0 (11.0)	14.7 (12.4)	n.r.	13.3 (6.6)	9.8 (4.0)	n.r.	p=0.006 [c]	0.07 [aa]	N/A
Henskens et al. (2018)	E-ADL Test		n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	N/A
Kampragkou et al. (2017)	One Leg Standing Balance Test [s]		4.0 (2.2)	3.5 (2.2)	n.r.	4.4 (3.1)	5.7 (3.0)	n.r.	F(1,26)=39.03, p<0.05 [i] CG: p>0.05 [l] IG: p=0.0001 [l]	n.r.	2.36
	TUG [s]		19.7 (4.4)	21.1 (4.8)	n.r.	19.3 (5.0)	18.2 (4.2)	n.r.	F(1,26)=22.09, p<0.05 [i] CG: p>0.05 [n] IG: p>0.05 [n]	n.r.	1.78
Kemoun et al. (2010)	Gait analysis	walking speed [m/s]	0.9 (0.2)	0.8 (0.2)	n.r.	0.7 (0.1)	1.0 (0.2)	n.r.	F(1,29)=53.4, p=0.01 [i]	n.r.	2.72
		stride length [m]	1.0 (0.2)	0.9 (0.2)	n.r.	0.9 (0.2)	1.0 (0.2)	n.r.	F(1,29)=16.3, p=0.01 [i]	n.r.	1.50
		double limb support time [s]	0.1 (0.0)	0.1 (0.0)	n.r.	0.2 (0.0)	0.1 (0.0)	n.r.	F(1,29)=27.0, p=0.01 [i]	n.r.	1.93
M.-J. Kim et al. (2016)	BBS		n.r.	n.r.	n.r.	28.2 (17.6)	21.5 (17.3)	n.r.	IG: p=0.04 [m]	n.r.	N/A
	Handgrip dynamometer [kg]		n.r.	n.r.	n.r.	7.9 (5.9)	11.8 (7.7)	n.r.	IG: p=0.02 [m]	n.r.	N/A
	Pedal Power	pedal rotation	n.r.	n.r.	n.r.	97.7 (89.9)	285.8 (197.5)	n.r.	IG: p=0.004 [m]	n.r.	N/A

		total load [W*number of pedal rotation/s]	n.r.	n.r.	n.r.	6.3 (7.5)	10.0 (6.8)	n.r.	IG: p=0.06 [m]	n.r.	N/A
Kovács et al. (2013)	POMA	total score	10 (n.r.)	11 (n.r.)	n.r.	14 (n.r.)	17 (n.r.)	n.r.	CG: p=0.624 [n] IG: p<0.0001 [n]	n.r.	N/A
		balance score	6 (n.r.)	7 (n.r.)	n.r.	7 (n.r.)	11 (n.r.)	n.r.	CG: p=0.640 [n] IG: p<0.0001 [n]	n.r.	N/A
		gait score	4 (n.r.)	4 (n.r.)	n.r.	5 (n.r.)	7 (n.r.)	n.r.	CG: p=0.530 [n] IG: p<0.0001 [n]	n.r.	N/A
	TUG [s]		32.1 (n.r.)	33.3 (n.r.)	n.r.	32.6 (n.r.)	31.1	n.r.	CG: p=0.171 [n] IG: p<0.0001 [n]	n.r.	N/A
Lam, Liao et al. (2018)	BBS		42.5 (10.7)	45.5 (10.6)	3.1 (n.r.)	43.6 (9.7)	45.2 (8.9)	1.6 (n.r.)	p=0.571 [o]	0.011 [z]	N/A
	POMA	total score	24.3 (5.8)	25.4 (5.5)	1.2 (n.r.)	25.4 (3.8)	26.0 (4.0)	0.6 (n.r.)	p=0.382 [o]	0.017 [z]	N/A
		balance score	14.1 (2.3)	14.0 (3.4)	0.7 (n.r.)	14.1 (2.3)	14.7 (2.4)	0.7 (n.r.)	p=0.705 [o]	0.006 [z]	N/A
		gait score	11.3 (1.6)	11.4 (2.3)	0.5 (n.r.)	11.3 (1.6)	11.3 (1.8)	-0.1 (n.r.)	p=0.178 [o]	0.034 [z]	N/A
	TUG [s]		23.0 (15.7)	21.3 (15.9)	-1.8 (n.r.)	20.3 (10.5)	19.8 (12.6)	-0.5 (n.r.)	p=0.707 [o]	0.006 [z]	N/A
	5x STS [s]		25.7 (16.6)	22.9 (11.2)	-2.8 (n.r.)	21.6 (8.2)	21.0 (8.4)	-0.6 (n.r.)	p=0.720 [o]	0.006 [z]	N/A
Miu et al. (2008)	FR [cm]		19.9 (8.3)	n.r.	n.r.	19.7 (7)	n.r.	n.r.	IG: p=0.007 [m]	n.r.	N/A
	BBS		n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	IG: p<0.001 [m]	n.r.	N/A
	6min WT [m]		264 (100)	n.r.	n.r.	245 (90)	n.r.	38 (42)	IG: p<0.001 [m]	n.r.	N/A

Netz et al. (2007)	FR [cm]		19.3 (10.3)	22.9 (7.0)	3.6 (SE: 3.7)	20.1 (7.9)	21.7 (6.9)	1.6 (SE: 1.3)	n.s. [i]	n.r.	0.27
	TUG [s]		16.5 (9.3)	14.7 (7.3)	-1.8 (SE: 0.8)	17.8 (8.4)	18.1 (8.8)	0.3 (SE: 0.7)	n.s. [i]	n.r.	0.85
	5x STS [s]		14.0 (4.1)	14.6 (4.1)	0.6 (SE: 1.0)	15.4 (4.2)	16.1 (5.8)	0.7 (SE: 1.0)	n.s. [i]	n.r.	0.03
Padala et al. (2017)	BBS		45.8 (2.5)	n.r.	n.r.	46.5 (2.4)	n.r.	n.r.	p=0.048 [i]	n.r.	N/A
Padala et al. (2012)	BBS		41.3 (7.6)	46.6 (8.7)	n.r.	43.4 (8.9)	49.6 (5.7)	n.r.	p=0.56 [i]	n.r.	N/A
	POMA		22.9 (2.6)	24.9 (3.4)	n.r.	23.5 (3.7)	25.3 (2.8)	n.r.	p=0.97 [i]	n.r.	N/A
	TUG [s]		14.9 (4.7)	12.8 (3.2)	n.r.	14.7 (7.2)	13.9 (7.9)	n.r.	p=0.52 [i]	n.r.	N/A
Pedrinolla et al. (2018)	Gait analysis	speed [cm/s]	92.3 (5.7)	n.r.	n.r.	92.5 (10.2)	n.r.	n.r.	z=-1.77, p=0.076 [p]	n.r.	N/A
		stride [cm]	104.3 (4.6)	n.r.	n.r.	111.3 (7.2)	n.r.	n.r.	z=-0.86, p=0.391 [p]	n.r.	N/A
		step [cm]	52.0 (2.3)	n.r.	n.r.	55.7 (3.6)	n.r.	n.r.	z=-0.72, p=0.471 [p]	n.r.	N/A
		single support [%]	36.1 (0.4)	n.r.	n.r.	37.2 (1.1)	n.r.	n.r.	z=0.00, p=1.000 [p]	n.r.	N/A
		double support [%]	27.5 (0.9)	n.r.	n.r.	26.2 (1.3)	n.r.	n.r.	z=0.85, p=0.394 [p]	n.r.	N/A
	3-speed walking test, VO_2 [ml/min*kg^{-1}]	speed 1	11.0 (0.9)	n.r.	n.r.	13.3 (0.6)	n.r.	n.r.	t=2.28, p=0.030 [a]	n.r.	0.81
		speed 2	12.0 (0.8)	n.r.	n.r.	12.0 (0.6)	n.r.	n.r.	t=2.94, p=0.006 [a]	n.r.	1.04
		speed 3	17.1 (1.0)	n.r.	n.r.	14.8 (0.7)	n.r.	n.r.	t=2.09, p=0.054 [a]	n.r.	0.74

	3-speed walking test, heart rate [bpm]	speed 1	93.0 (2.8)	n.r.	n.r.	98.3 (2.3)	n.r.	n.r.	t=1.58, p=0.126 [a]	n.r.	0.56
		speed 2	99.2 (2.8)	n.r.	n.r.	103.9 (3.0)	n.r.	n.r.	t=2.72, p=0.011 [a]	n.r.	0.96
		speed 3	106.8 (3.2)	n.r.	n.r.	108.2 (4.6)	n.r.	n.r.	t=1.75, p=0.107 [a]	n.r.	0.62
	3-speed walking test, energy cost of walking [J/kg*m^{-1}]	speed 1	6.5 (2.6)	n.r.	n.r.	6.2 (1.1)	n.r.	n.r.	z=2.04, p=0.041 [p]	n.r.	N/A
		speed 2	4.6 (1.0)	n.r.	n.r.	5.1 (0.5)	n.r.	n.r.	z=2.96, p=0.003 [p]	n.r.	N/A
		speed 3	5.9 (0.9)	n.r.	n.r.	6.2 (1.6)	n.r.	n.r.	z=1.47, p=0.142 [p]	n.r.	N/A
Pitkälä, Pöysti et al. (2013)	SPPB		9.7 (2.1)	n.r.	n.r.	IG1: 9.3 (2.4) IG2: 9.8 (2.2)	n.r.	n.r.	p=0.90 [q]	n.r.	N/A
Pomeroy et al. (1999)	Southampton Assessment of Mobility		13.3 (6.0)	12.2 (6.4)	-1.1 (4.1)	15.4 (4.1)	15.1 (4.2)	-0.4 (2.2)	p=0.614 [b]	n.r.	0.22
	2-min walk test [m]		23.8 (22.1)	24.4 (20.4)	n.r.	32.2 (15.7)	35.3 (18.6)	3.1 (9.3)	p=0.325 [a]	n.r.	N/A
Roach et al. (2011)	Acute Care Index of Function	transfer score	0.8 (0.2)	0.8 (0.3)	n.r.	IG1: 0.8 (0.2) IG2: 0.9 (0.2)	IG1: 0.9 (0.2) IG2: 0.8 (0.2)	n.r.	p=0.04 [i]	n.r.	N/A
		bed mobility score	0.9 (0.3)	0.8 (0.3)	n.r.	IG1: 0.9 (0.2) IG2: 0.8 (0.3)	IG1: 0.9 (0.2) IG2: 0.8 (0.3)	n.r.	p=0.77 [i]	n.r.	N/A
	6min WT [ft]		296.6 (229.4)	324.8 (274.4)	n.r.	IG1: 387.1 (214.8) IG2: 329.9 (247.4)	IG1: 384.9 (217.6) IG2: 367.5 (300.2)	n.r.	p=0.61 [i]	n.r.	N/A

Rolland et al. (2007)	One-leg balance test [% abnormal score]		(92.5 %)	(98.1 %)	n.r.	(91.0 %)	(94.6 %)	n.r.	p=0.34 [r]	n.r.	N/A
	Get-Up and Go Test		2.7 (0.8)	3.2 (1.2)	n.r.	2.7 (0.8)	3.1 (1.1)	n.r.	p=0.31 [r]	n.r.	N/A
	6m WT [m/s]		0.3 (0.1)	0.4 (0.2)	n.r.	0.3 (0.1)	0.4 (0.2)	n.r.	p=0.002 [r]	n.r.	N/A
Santana-Sosa et al. (2008)	POMA		n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	F(1,14)=45.13, p<0.001 [i]	0.887 [z]	3.59
	Senior Fitness test	30s CST	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	F(1,14)=48.74, p<0.001 [i]	0.777 [z]	3.73
		Arm curl test	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	F(1,14)=73.15, p<0.001 [i]	0.839 [z]	N/A
		Chair sit-and-reach test [cm]	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	F(1,14)=40.18, p<0.001 [i]	0.742 [z]	N/A
		Back scratch test [cm]	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	F(1,14)=36.04, p<0.001 [i]	0.720 [z]	N/A
		8-foot up-and-go test [s]	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	F(1,14)=36.78, p<0.001 [i]	0.724 [z]	N/A
		2-min step test	n.r.	n.r.	n.r.	n.r.	n.r.	n.r.	F(1,14)=8.96, p=0.010 [i]	0.390 [z]	N/A
Schwenk, Dutzi et al. (2014)	Postural sway (inertial sensors) [sq cm]		6.6 (0.8)	6.5 (0.8)	n.r.	6.5 (0.6)	6.1 (0.6)	n.r.	p=0.023 [i]	0.06 [z]	N/A
	Gait analysis	speed [cm/sec]	73.3 (37.1)	89.4 (36.8)	n.r.	72.7 (38.6)	92.8 (37.6)	n.r.	p=0.354 [i]	0.01 [z]	N/A
		stride length [cm]	80.9 (28.9)	91.2 (27.8)	n.r.	83.2 (31.0)	96.4 (29.9)	n.r.	p=0.354 [i]	0.01 [z]	N/A
		cadence [steps/min]	105.8 (22.0)	115.7 (19.8)	n.r.	100.6 (25.7)	113.5 (24.6)	n.r.	p=0.343 [i]	0.01 [z]	N/A
	Hierarchical Assessment of Balance and Mobility		40.9 (13.6)	46.4 (13.0)	n.r.	38.5 (15.1)	46.6 (10.2)	n.r.	p=0.162 [i]	0.02 [z]	N/A

	5x STS [s]		16.4 (6.8)	15.8 (10.9)	n.r.	17.0 (6.4)	13.1 (4.6)	n.r.	p=0.037 [i]	0.06 [z]	N/A
	Handgrip dynamometer [kg]		14.6 (6.2)	15.1 (6.6)	n.r.	14.4 (6.2)	14.8 (6.7)	n.r.	p=0.834 [i]	0.00 [z]	N/A
	One-repetition maximum in leg press [kg]		97.0 (51.0)	102.2 (54.4)	n.r.	99.7 (59.4)	140.0 (70.2)	n.r.	p<0.001 [i]	0.36 [z]	N/A
	One-repetition maximum of abductor [kg]		66.9 (28.8)	69.7 (29.0)	n.r.	70.8 (34.9)	88.2 (36.8)	n.r.	p<0.001 [i]	0.11 [z]	N/A
Schwenk, Zieschang et al. (2014)	Gait analysis	speed [cm/sec]	128.7 (38.2)	127.6 (35.7)	n.r.	132.7 (55.7)	149.3 (48.2)	n.r.	p<0.001 [c]	1.27 [y]	N/A
		cadence [steps/min]	134.5 (17.9)	132.0 (19.2)	n.r.	137.1 (21.1)	145.4 (20.8)	n.r.	p=0.002 [c]	0.96 [y]	N/A
		stride length [cm]	115.3 (29.5)	115.9 (25.7)	n.r.	116.6 (42.6)	124.8 (37.4)	n.r.	p=0.008 [c]	0.80 [y]	N/A
		stride time [sec]	0.9 (0.1)	0.9 (0.1)	n.r.	0.9 (0.2)	0.8 (0.1)	n.r.	p=0.001 [c]	0.99 [y]	N/A
		double support [%]	25.9 (6.1)	25.4 (6.0)	n.r.	26.9 (8.9)	23.0 (7.8)	n.r.	p=0.001 [c]	1.03 [y]	N/A
		step width [cm]	10.2 (4.2)	9.9 (4.4)	n.r.	11.3 (4.2)	11.1 (5.0)	n.r.	p=0.999 [c]	0.00 [y]	N/A
		step time variability [CV]	5.0 (2.5)	5.4 (2.6)	n.r.	5.2 (3.4)	5.1 (2.1)	n.r.	p=0.425 [c]	0.22 [y]	N/A
		Walk-Ratio	0.4 (0.1)	0.4 (0.1)	n.r.	0.4 (0.2)	0.4 (0.2)	n.r.	p=0.554 [c]	0.18 [y]	N/A
Schwenk et al. (2010)	Gait analysis, dual task cost,	gait speed	-22.6 (18.4)	-20.8 (15.8)	n.r.	-21.9 (11.9)	-13.5 (9.4)	n.r.	p=0.086 [i]	n.r.	N/A
		cadence	-18.8 (15.4)	-14.9 (12.5)	n.r.	-17.5 (10.4)	-12.8 (10.1)	n.r.	p=0.846 [i]	n.r.	N/A

	serial 2 condition [%]	stride length	-5.7 (12.3)	-6.8 (13.6)	n.r.	-5.3 (9.2)	-0.1 (8.2)	n.r.	p=0.074 [i]	n.r.	N/A
		stride time	30.4 (45.8)	21.0 (23.0)	n.r.	24.1 (20.0)	16.8 (17.7)	n.r.	p=0.750 [i]	n.r.	N/A
		single support	-6.0 (10.7)	-6.2 (7.0)	n.r.	-4.7 (6.4)	-3.2 (4.2)	n.r.	p=0.459 [i]	n.r.	N/A
		motor + cognitive performance	-14.7 (21.3)	-13.6 (17.1)	n.r.	-18.1 (15.2)	-12.3 (7.8)	n.r.	p=0.378 [i]	n.r.	N/A
	Gait analysis, dual task cost, serial 3 condition [%]	gait speed	-39.8 (18.9)	-37.2 (16.7)	n.r.	-41.6 (18.4)	-20.0 (12.7)	n.r.	p<0.001 [i]	n.r.	N/A
		cadence	-26.8 (15.8)	-23.6 (14.1)	n.r.	-27.9 (18.5)	-15.3 (11.0)	n.r.	p=0.007 [i]	n.r.	N/A
		stride length	-18.8 (14.0)	-18.0 (15.6)	n.r.	-20.7 (12.2)	-5.6 (11.7)	n.r.	p=0.001 [i]	n.r.	N/A
		stride time	44.0 (35.8)	35.8 (27.8)	n.r.	62.0 (102.0)	20.9 (20.7)	n.r.	p=0.056 [i]	n.r.	N/A
		single support	-9.7 (11.3)	-10.6 (8.7)	n.r.	-13.8 (13.4)	-5.1 (5.0)	n.r.	p=0.003 [i]	n.r.	N/A
		motor + cognitive performance	-31.9 (20.3)	-29.7 (17.5)	n.r.	-32.7 (24.8)	-12.1 (18.4)	n.r.	p=0.026 [i]	n.r.	N/A
Sobol et al. (2016)	TUG [s]		6.6 (1.62)	6.6 (1.9)	n.r.	6.7 (1.8)	6.5 (1.7)	n.r.	p=0.151 [s]	n.r.	N/A
	10-metre walk test [m/s]	single task	1.4 (0.3)	1.4 (0.2)	n.r.	1.3 (0.2)	1.3 (0.2)	n.r.	p=0.108 [s]	n.r.	N/A
		dual task months	1.0 (0.3)	1.1 (0.3)	n.r.	0.9 (0.3)	1.0 (0.4)	n.r.	p=0.051 [s]	n.r.	N/A
		dual task numbers	1.1 (0.3)	1.1 (0.3)	n.r.	0.9 (0.3)	1.0 (0.4)	n.r.	p=0.155 [s]	n.r.	N/A

	30s CST		14.9 (4.2)	15.5 (4.3)	n.r.	13.9 (3.6)	14.3 (3.5)	n.r.	p=0.408 [s]	n.r.	N/A
	400-m walk test [s]		305 (71.5)	303 (70.5)	n.r.	306 (92.3)	296 (79.1)	n.r.	p=0.118 [s]	n.r.	N/A
	6-min Astrand Cycle Ergometer test, VO_{2max} [mL/kg/min]		26.2 (9.0)	27.2 (8.7)	n.r.	25.3 (7.5)	30.1 (7.5)	n.r.	p<0.0001 [s]	n.r.	N/A
Souto Barreto et al. (2017)	4m WT [m/s]		0.5 (0.2)	n.r.	0.03 (SE: 0.03)	0.5 (0.2)	n.r.	0.07 (SE: 0.03)	β=0.01, p=0.30 [t]	n.r.	0.20
	SPPB		4.5 (2.3)	n.r.	-0.8 (SE: 0.34)	4.4 (2.4)	n.r.	-0.2 (SE: 0.36)	β=0.10, p=0.22 [t]	n.r.	0.26
Steinberg et al. (2009)	8-foot walk test [s]		3.7 (1.6)	n.r.	n.r.	3.6 (1.8)	n.r.	n.r.	β=-0.08 (0.27), p=0.77 [u]	n.r.	N/A
	5x STS [s]		16.1 (6.5)	n.r.	n.r.	16.8 (7.4)	n.r.	n.r.	β=-4.4 (3.6), p=0.22 [u]	n.r.	N/A
	Jebsen Total Time [s]		107.3 (49.9)	n.r.	n.r.	83.5 (41.9)	n.r.	n.r.	β=-23.39 (11.6), p=0.04 [u]	n.r.	N/A
Suttanon et al. (2013)	Modified Clinical Test of Sensory Interaction of Balance [deg/s]		1.5 (0.7)	1.7 (0.8)	n.r.	1.9 (0.7)	1.8 (0.7)	n.r.	p=0.086 [s]	n.r.	N/A
	Limits of stability	reaction time [ms]	1.1 (0.2)	1.2 (0.3)	n.r.	1.2 (0.3)	1.1 (0.2)	n.r.	p=0.365 [s]	n.r.	N/A
		movement velocity [degrees/s]	3.1 (1.2)	3.4 (1.0)	n.r.	3.0 (1.3)	3.0 (1.1)	n.r.	p=0.016 [s]	n.r.	N/A
		maximum excursion [%]	72.4 (12.0)	72.7 (12.1)	n.r.	66.3 (14.4)	68.3 (15.5)	n.r.	p=0.817 [s]	n.r.	N/A

	directional control [%]	64.4 (10.0)	61.3 (11.0)	n.r.	60.3 (12.3)	60.7 (11.3)	n.r.	p=0.446 [s]	n.r.	N/A
FR [cm]		28.5 (4.7)	25.5 (5.3)	n.r.	23.5 (5.7)	25.8 (5.6)	n.r.	p=0.002 [s]	n.r.	N/A
Hill Step Test		13.0 (3.2)	11.8 (3.5)	n.r.	12.3 (2.4)	12.3 (3.0)	n.r.	p=0.082 [s]	n.r.	N/A
Step quick turn, worse side	time [s]	3.3 (1.0)	3.1 (1.1)	n.r.	3.8 (1.7)	3.7 (2.0)	n.r.	p=0.283 [s]	n.r.	N/A
	sway [deg]	48.9 (8.2)	47.3 (6.7)	n.r.	49.0 (11.1)	48.5 (13.0)	n.r.	p=0.452 [s]	n.r.	N/A
TUG [s]		16.4 (6.6)	16.6 (6.2)	n.r.	16.2 (5.0)	16.2 (5.6)	n.r.	p=0.571 [s]	n.r.	N/A
Cognitive TUG [s]		18.1 (3.4)	19.2 (6.0)	n.r.	25.4 (8.0)	23.2 (7.7)	n.r.	p=0.994 [s]	n.r.	N/A
Manual TUG [s]		18.0 (6.8)	19.0 (7.3)	n.r.	18.4 (5.8)	18.2 (6.6)	n.r.	p=0.088 [s]	n.r.	N/A
Gait analy-sis	step width [cm]	15.6 (4.5)	16.2 (4.0)	n.r.	16.2 (2.3)	15.6 (2.5)	n.r.	p=0.125 [s]	n.r.	N/A
	step length [cm]	36.8 (13.2)	36.0 (9.5)	n.r.	32.5 (8.3)	31.8 (10.7)	n.r.	p=0.907 [s]	n.r.	N/A
	speed [cm/s]	40.4 (13.5)	41.7 (14.3)	n.r.	39.4 (11.6)	38.9 (13.6)	n.r.	p=0.244 [s]	n.r.	N/A
5x STS [s]		13.3 (5.0)	13.3 (3.7)	n.r.	13.2 (4.2)	14.6 (5.1)	n.r.	p=0.945 [s]	n.r.	N/A
STS on NeuroCom Balance Master	Raising index [% body weight]	16.3 (4.8)	17.0 (6.9)	n.r.	13.5 (4.7)	14.5 (6.1)	n.r.	p=0.725 [s]	n.r.	N/A
	sway [deg/s]	4.2 (1.3)	4.7 (1.5)	n.r.	4.0 (1.1)	4.3 (1.1)	n.r.	p=0.290 [s]	n.r.	N/A
Physiological Profile Assessment [score]		1.4 (1.2)	1.8 (1.2)	n.r.	1.8 (1.2)	1.9 (0.3)	n.r.	p=0.314 [s]	n.r.	N/A

Tappen et al. (2000)	6min WT [ft]		261.1 (175.0)	212.1 (168.8)	n.r.	IG1: 330.2 (250.0) IG2: 391.7 (233.3)	IG1: 321.9 (223.2) IG2: 310.6 (219.3)	n.r.	p<0.05 [i] CG: p=0.0874 [e] IG1: n.s. [e] IG2: p=0.0119 [e] F=5.59, p<0.01 [v]	n.r.	N/A
Telenius et al. (2015a)	BBS [score]		35.4 (13.7)	36.6 (14.4)	n.r.	34.3 (14.5)	37.2 (14.0)	n.r.	p=0.02 [a]	0.4 [y]	N/A
	6m WT [m/s]		0.5 (0.2)	0.5 (0.3)	n.r.	0.5 (0.2)	0.5 (0.2)	n.r.	p=0.86 [a]	0.0 [y]	N/A
	30s CST		6.2 (2.9)	6.6 (3.7)	n.r.	6.0 (3.1)	7 (3.3)	n.r.	p=0.11 [a]	0.2 [y]	N/A
Toots et al. (2017)	4m WT [m/s]	walking aid	0.5 (0.2)	n.r.	-0.02 (SE: 0.02)	0.5 (0.2)	n.r.	-0.02 (SE: 0.02)	p=0.777 [w]	-0.05 [bb]	0.05
		no walking aid	0.5 (0.2)	n.r.	-0.02 (SE: 0.02)	0.4 (0.2)	n.r.	0.01 (SE: 0.02)	p=0.242 [w]	0.20 [bb]	0.20
Toots et al. (2016)	BBS		29.3 (14.7)	n.r.	-1.8 (SE: 0.9)	28.6 (14.3)	n.r.	2.4 (SE: 0.9)	p<0.001 [w]	0.52 [bb]	0.53
Toulotte et al. (2003)	Postural sway (posturography platform) [mm²]		292.3 (94.5)	n.r.	n.r.	398.7 (229.6)	n.r.	n.r.	p<0.01 [i]	n.r.	N/A
	TUG [s]		39.4 (17.7)	n.r.	n.r.	67.6 (38.9)	n.r.	n.r.	p<0.01 [i]	n.r.	N/A
	10-metre walk test [s]		63.4 (51.1)	n.r.	n.r.	60.6 (49.9)	n.r.	n.r.	p<0.05 [i]	n.r.	N/A
	Chair sit-and-reach test [cm]		10.3 (8.0)	n.r.	n.r.	11.4 (7.1)	n.r.	n.r.	p<0.05 [i]	n.r.	N/A
Venturelli et al. (2011)	6min WT [m]		238 (47)	168 (34)	n.r.	245 (31)	294 (49)	n.r.	p<0.001 [i]	n.r.	N/A

Vreugdenhil et al. (2012)	FR [cm]		24.0 (6.4)	22.1 (7.9)	-1.9 (SE: 1.3)	27.6 (7.4)	30.6 (7.0)	2.3 (SE: 1.1)	p=0.032 [s]	n.r.	0.80
	TUG [s]		11.1 (3.3)	12.8 (4.1)	2.0 (SE: 0.7)	9.7 (3.7)	9.1 (3.8)	-0.9 (SE: 0.5)	p=0.004 [s]	n.r.	1.09
	10-s chair-stand test		8.5 (2.9)	7.2 (3.2)	-1.0 (SE: 0.4)	9.2 (2.5)	10.8 (2.0)	1.7 (SE: 0.4)	p<0.001 [s]	n.r.	1.55
Werner et al. (2017)	Body-fixed-sensor-based STS analysis (DynaPort)	trunk flexion, range [°]	33.1 (9.9)	34.9 (10.5)	0.9 (9.0)	33.6 (7.5)	40.2 (12.8)	8.3 (13.4)	p=0.006 [i]	0.099 [z]	0.66
		trunk flexion, duration [s]	1.1 (0.5)	1.1 (0.5)	-0.1 (0.6)	1.2 (0.9)	1.8 (1.1)	0.8 (1.2)	p<0.001 [i]	0.188 [z]	0.96
		maximum trunk flexion, angular velocity [°/s]	73.9 (26.6)	81.4 (26.8)	5.9 (22.0)	79.0 (27.4)	64.8 (27.1)	-12.3 (26.4)	p=0.002 [i]	0.127 [z]	0.76
		STS movement duration [s]	2.1 (0.9)	2.0 (0.7)	-0.1 (0.9)	2.1 (1.1)	2.8 (1.4)	0.9 (1.5)	p<0.001 [i]	0.158 [z]	0.87
	ACSID	recall and initiation score	1.8 (1.0)	1.7 (.9)	-0.1 (0.9)	1.8 (1.0)	3.3 (1.3)	1.7 (1.7)	p<0.001 [i]	0.319 [z]	1.35
		effective performance score	1.9 (.9)	1.8 (.8)	-0.1 (0.9)	2.1 (.9)	3.1 (1.0)	1.0 (0.9)	p<0.001 [i]	0.261 [z]	1.24
		total score	3.7 (1.4)	3.5 (1.5)	-0.1 (1.4)	3.9 (1.4)	6.4 (2.1)	2.7 (2.2)	p<0.001 [i]	0.372 [z]	1.54
Wesson et al. (2013)	Near-tandem test [n.r.]		5.7 (3.0)	6.3 (3.7)	n.r.	5.2 (3.6)	5.4 (3.7)	n.r.	p=0.32 [b]	n.r.	N/A
	Hill Step Test		14.5 (5.0)	14.2 (7.7)	n.r.	19.2 (6.5)	15.0 (5.1)	n.r.	p=0.1 [b]	n.r.	N/A

	Physiological Profile Assessment		1.7 (1.7)	2.7 (1.8)	n.r.	0.8 (1.2)	1.4 (1.6)	n.r.	p=0.82 [b]	n.r.	N/A
Wiloth et al. (2018)	Physiomat-Follow-The-Ball Task, trained	duration [s]	28.9 (15.7)	23.4 (5.5)	n.r.	30.9 (17.5)	19.3 (4.6)	n.r.	p<0.001 [c]	0.253 [aa]	N/A
		accuracy [digits/ms]	4164.3 (3922.4)	3776.3 (1286.9)	n.r.	4450.4 (2859.8)	3169.7 (557.2)	n.r.	p<0.001 [c]	0.144 [aa]	N/A
	Physiomat-Trail-Making Task, Level 1 trained	duration [s]	11.5 (4.7)	9.7 (3.1)	n.r.	16.7 (20.3)	7.2 (1.9)	n.r.	p<0.001 [c]	0.260 [aa]	N/A
		accuracy [digits/ms]	2108.1 (911.9)	2124.9 (773.9)	n.r.	2849.1 (4199.2)	1782.0 (339.7)	n.r.	p=0.007 [c]	0.092 [aa]	N/A
	Physiomat-Trail-Making Task, Level 2 trained	duration [s]	19.9 (11.5)	16.7 (4.9)	n.r.	21.8 (9.7)	14.3 (5.6)	n.r.	p<0.001 [c]	0.311 [aa]	N/A
		accuracy [digits/ms]	3005.7 (1066.2)	3187.0 (940.1)	n.r.	3390.1 (1800.1)	2923.1 (803.0)	n.r.	p=0.003 [c]	0.127 [aa]	N/A
	Physiomat-Trail-Making Task, Level 3 trained	duration [s]	25.9 (10.4)	22.7 (5.3)	n.r.	28.6 (11.8)	20.0 (7.4)	n.r.	p<0.001 [c]	0.293 [aa]	N/A
		accuracy [digits/ms]	3742.8 (557.8)	3992.0 (945.8)	n.r.	4376.7 (1528.5)	3806.1 (1246.3)	n.r.	p=0.047 [c]	0.065 [aa]	N/A
	Physiomat-Trail-Making Task, Level 4 trained	duration [s]	43.9 (9.0)	44.0 (15.5)	n.r.	51.3 (16.6)	34.9 (7.9)	n.r.	p<0.001 [c]	0.340 [aa]	N/A
		accuracy [digits/ms]	7724.4 (1676.4)	7880.0 (2238.8)	n.r.	8176.3 (2484.2)	6599.3 (1468.6)	n.r.	p<0.001 [c]	0.365 [aa]	N/A
	Physiomat-Trail-Making Task, Level 5 trained	duration [s]	58.7 (17.6)	56.2 (16.7)	n.r.	56.3 (12.1)	48.7 (14.7)	n.r.	p<0.001 [c]	0.589 [aa]	N/A
		accuracy [digits/ms]	8467.8 (1646.8)	9360.2 (2855.1)	n.r.	8444.3 (2261.5)	8005.0 (1906.4)	n.r.	p=0.007 [c]	0.329 [aa]	N/A
	Physiomat-Trail-Making Task score, trained		3.2 (1.5)	3.6 (1.4)	n.r.	3.4 (1.4)	4.7 (.9)	n.r.	p<0.001 [c]	0.211 [aa]	N/A

	Physiomat-Trail-Making Task, Level 1 untrained	duration [s]	11.2 (5.4)	9.7 (3.1)	n.r.	15.6 (15.6)	7.5 (2.1)	n.r.	p<0.001 [c]	0.219 [aa]	N/A
		accuracy [digits/ms]	2043.6 (1246.8)	1959.8 (543.6)	n.r.	2523.1 (3040.6)	1735.4 (317.1)	n.r.	p=0.017 [c]	0.073 [aa]	N/A
	Physiomat-Trail-Making Task, Level 2 untrained	duration [s]	17.9 (8.4)	14.7 (3.7)	n.r.	18.0 (7.4)	13.9 (9.9)	n.r.	p<0.001 [c]	0.236 [aa]	N/A
		accuracy [digits/ms]	2770.9 (1598.6)	2661.6 (785.8)	n.r.	2703.4 (1586.1)	2683.1 (1746.5)	n.r.	p=0.121 [c]	0.037 [aa]	N/A
	Physiomat-Trail-Making Task, Level 3 untrained	duration [s]	29.6 (15.9)	25.4 (7.0)	n.r.	32.9 (16.5)	23.1 (10.9)	n.r.	p<0.001 [c]	0.204 [aa]	N/A
		accuracy [digits/ms]	4467.0 (1093.5)	4816.0 (1671.0)	n.r.	5316.9 (2637.1)	4539.6 (2924.3)	n.r.	p=0.008 [c]	0.122 [aa]	N/A
	Physiomat-Trail-Making Task, Level 4 untrained	duration [s]	41.4 (9.5)	38.7 (12.2)	n.r.	38.7 (11.8)	33.0 (7.7)	n.r.	p=0.005 [c]	0.280 [aa]	N/A
		accuracy [digits/ms]	8137.4 (2374.7)	7671.3 (3642.1)	n.r.	7285.2 (2346.8)	5857.7 (1296.5)	n.r.	p=0.009 [c]	0.244 [aa]	N/A
	Physiomat-Trail-Making Task, Level 5 untrained	duration [s]	55.5 (11.8)	54.5 (15.1)	n.r.	55.7 (16.5)	48.5 (13.7)	n.r.	p=0.003 [c]	0.384 [aa]	N/A
		accuracy [digits/ms]	8598.2 (1539.8)	9027.3 (2444.4)	n.r.	9227.8 (4138.9)	8278.5 (2000.1)	n.r.	P00.001 [c]	0.459 [aa]	N/A
	Physiomat-Trail-Making Task score, untrained		3.0 (1.5)	3.5 (1.4)	n.r.	3.1 (1.4)	4.5 (1.0)	n.r.	p<0.001 [c]	0.184 [aa]	N/A
Yoon et al. (2013)	Postural sway (Wii Balance Board), eyes closed	wide base, COP velocity [cm/s]	3.5 (1.7)	3.2 (1.3)	0.8 (1.8)	4.8 (1.4)	3.6 (1.2)	1.2 (1.0)	p<0.05 [b] IG: p<0.05 [x]	n.r.	0.30
		narrow base, COP path length [cm/s]	104.7 (46.6)	95.4 (37.5)	9.3 (14.1)	146.4 (64.5)	93.9 (38.9)	52.5 (59.7)	p<0.05 [b] IG: p<0.05 [x]	n.r.	1.00

BBS	34.9 (4.6)	35.1 (4.4)	-0.2 (2.4)	35.3 (1.8)	38.0 (2.0)	-2.7 (0.9)	p<0.05 [b] CG / IG: p<0.05 [x]	n.r.	1.52
TUG [s]	28.8 (5.7)	n.r.	n.r.	27.7 (6.1)	n.r.	n.r.	n.r.	n.r.	N/A

4m WT: 4-metre walk test, 5x STS: Five Times Sit-to-Stand Test, 6m WT: 6-metre walk test, 6min WT: 6-minute walk test, 30s CST: 30-second chair stand test, ACSID: Assessment of Compensatory Sit-to-Stand Maneuvers in People With Dementia, BBS: Berg Balance Scale, bpm: beats per minute, CG: control group, COP: centre of pressure, E-ADL Test: Erlangen Test of Activities of Daily Living, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, FR: Functional Reach Test, GMWT: Groningen Meander Walking Test, IG: intervention group, N/A: not applicable, n.s.: not significant, POMA: Performance Oriented Mobility Assessment, SD: standard deviation, SE: standard error, SPPB: Short Physical Performance Battery, STS: Sit-to-Stand, TUG: Timed Up & Go Test, VO_2(max): (maximum) oxygen uptake

[a] independent t-test, between-group baseline-post difference, [b] Mann-Whitney test, between-group baseline-post difference, [c] Analysis of covariance with baseline scores as covariates, between-group post difference, [d] Kruskal-Wallis tests, between-group baseline-post difference, [e] dependent t-test, within-group baseline-post difference, [f] independent t-test, between-group post difference, [g] mixed model analysis of covariance with repeated measures, [h] general linear model analysis of covariance with repeated-measures, [i] two-way analysis of variance with repeated measures, group*time interaction, [j] one-way analysis of variance, between-group post difference, [k] linear regression analyses, between-group baseline-post difference, [l] Tukey test, within-group baseline-post difference, [m] one-way analysis of variance/analysis of variance with repeated measures, within-group baseline-post difference, [n] Friedman analysis of variance, within-group baseline-post difference, [o] mixed design multivariate analysis of variance, group*time interaction, [p] Wilcoxon's test, between-group baseline-post difference, [q] Analysis of covariance with age, sex, and use of mobility devices, between-group baseline-post difference, [r] ?, between-group baseline-post difference, [s] general linear model analysis, between-group baseline-post difference, [t] three level regression model, between-group baseline-post difference, [u] linear random effects models, time*group interaction, [v] Analysis of covariance with baseline scores and treatment fidelity as covariates, between-group post difference, [w] linear mixed effects models between-group baseline-post difference, [x] Wilcoxon's signed-rank test, within-group baseline-post difference, [y] standardized mean difference/Cohen's d, between-group, [z] partial eta squared, time*group interaction effect, [aa] partial eta squared, group effect, [bb] between-group difference/unadjusted pooled standard deviation of baseline-post difference

* Intention-to-treat analysis and complete-case analysis

Additional file 8

Data extraction main search

Content, construct, and criterion validity, internal consistency, and intra-rater reliability

***Table D.** Data extraction main search - content, construct, and criterion validity, internal consistency, and intra-rater reliability

Motor Assessment	Outcome	Reference
	Construct validity	
Physiomat-Trail-Making Task	Hypotheses testing 1) moderate-to-high associations with MMSE: r=0.29-0.66, p≤0.001-0.004 → Yes 2) more pronounced associations with modified Trail-Making-Test A: r=0.36-0.83, p≤0.001 → Yes 3) moderate associations with memory tests: r=(-0.42)-(-0.16), p=0.004-0.12 → Yes 4) higher associations of cognitive outcome measures with increasing complexity: r=(-0.33)-0.36, p≤0.001-0.02 vs. r=(-0.42)-0.44, p≤0.001-0.12 vs. r=(-0.22)-0.83, p≤0.001-0.12 → Yes (repeating numbers) 5) associations with TUG and POMA: r=(-0.40)-0.48, p≤0.001-0.71 → Yes 6) pronounced associations with Physiomat-Follow-The-Ball Task: r=0.61-0.71, p≤0.001 → Yes 7) less association with moderate Physiomat-Balance-Task (10 seconds): r=(-0.34)-0.11, p=0.10-0.71 → Yes 8) higher associations of motor-functional outcomes with decreasing complexity: r=(-0.22)-0.22, p=0.004-0.03 vs. r=(-0.40)-0.48, p≤0.001 vs. r=0.08-0.19, p=0.35-0.71 → No	Wiloth et al. (2016)
Maximum isometric strength assessed with dynamometers (knee extensor strength)	Mann-Whitney U-test/unpaired t-test, independent gait/STS performance vs. dependent gait/STS performance: p<0.0001	Suzuki et al. (2009)
E-ADL Test	Hypotheses testing Correlation with severity of dementia: r=(-0.47)-0.72, p≤0.001 Correlation with Nurses' Observations Scale for Geriatric Patients - instrumental activities of daily living/activities of daily living: r=(-0.45)-(-0.33), p≤0.001-0.023	Graessel et al. (2009)

	Correlation with Nurses' Observations Scale for Geriatric Patients - mood/disturbing behaviour: r=(-0.40)-(-0.33), p=0.007-0.027 Correlation with Nurses' Observations Scale for Geriatric Patients - total score: r=-0.60, p<0.001	
	Hypotheses testing Spearman correlation with cognition: r=0.39-0.43 Spearman correlation with everyday practical capabilities: r=0.39-0.64 Spearman correlation with mood/behaviour: r=0.11-0.39	Luttenberger et al. (2012)
Criterion validity (concurrent and predictive validity)/correlation with/prediction of external criteria		
Modified BBS	Bivariate correlation with spatiotemporal gait parameters: r=(-0.85)-0.73, n.s./p<0.05/p<0.01 Partial correlation with spatiotemporal gait parameters: r=(-0.67)-0.72, n.s./p<0.05/p<0.01	McGough et al. (2013)
POMA	Mann-Whitney U-test, fallers vs. non-fallers: p=0.928 Univariate logistic regression analysis to predict risk of falling in the next three months: R^2=0.000, OR=1.002, CI_{95}=0.904-1.111, p=0.966	Schwenk, Hauer et al. (2014)
TUG	Mann-Whitney U-test: fallers vs. non-fallers: p=0.236 Univariate logistic regression analysis to predict risk of falling in the next three months: R^2=0.011, OR=0.966, CI_{95}=0.883-1.056, p=0.612	Schwenk, Hauer et al. (2014)
5x STS	Mann-Whitney U-test, fallers vs. non-fallers: p=0.553 Univariate logistic regression analysis to predict risk of falling in the next three months: R^2=0.005, OR=1.023, CI_{95}=0.937-1.118, p=0.966	Schwenk, Hauer et al. (2014)
ACSID	Correlation with 2D video-motion analysis: r=(-0.73)-0.84, p<0.001	Werner et al. (2018)
Maximum isometric strength assessed with dynamometers (knee extensor strength)	Logistic regression analysis: knee extensor strength was a significant predictor of - Gait performance (OR: 443.02, CI_{95}: 9.20-21325.69) - STS performance (OR: 47.32, CI_{95}: 3.31-675.81)	Suzuki et al. (2009)
	Chi^2 test/Mann-Whitney U-test/ unpaired t-test, independent activities of daily living/gait performance vs. dependent activities of daily living/gait performance: p≤0.0001 Logistic regression analysis: knee extensor strength muscles was a significant predictor of - Dressing the lower body (OR: 109.90, CI_{95}: 7.60-1589.49) - Toileting (OR: 18.29, CI_{95}: 2.41-138.84) - Transfers to bed/toilet/shower (OR: 39.70, CI_{95}: 4.51-349.08) - Gait performance (OR: 12.77, CI_{95}: 2.30-70.77)	Suzuki et al. (2012)

6min WT	Pearson bivariate correlation with peak cycle ergometer test: r=0.33-0.51, p<0.05	Bronas et al. (2017)
SPPB	Bivariate correlation with spatiotemporal gait parameters: r=(-0.71)-0.66, n.s./p<0.01 Partial correlation with spatiotemporal gait parameters: r=(-0.65)-0.71, n.s./p<0.05/p<0.01	McGough et al. (2013)
	Pearson bivariate correlation with peak cycle ergometer test: r=0.35, p<0.05	Bronas et al. (2017)
E-ADL Test	Correlation with level of care: eta=0.39 Degree of level of care in relation to E-ADL score: eta=0.48 Kruskal-Wallis test: p<0.001 (df=2) for care level at baseline and after 22 months (df=3) Mann-Whitney U-test, unchanged care level vs. increased care level: p=0.01, U=376, achieved power at p=0.01: 0.48	Luttenberger et al. (2012)
	Internal consistency	
BBS	Cronbach's α=0.95 Item-to-total correlations: r>0.4 for all items except item 3, n.s./p<0.05/p<0.01	Telenius et al. (2015b)
E-ADL Test	Cronbach's α=0.77 Correlation between items: r=0.18-0.51, p<0.001-0.224	Graessel et al. (2009)
	Cronbach's α=0.68 (total sample), α=0.37 (mild dementia), α=0.64 (moderate dementia), α=0.73 (severe dementia) Correlation between the 5 items: r=0.21-0.44	Luttenberger et al. (2012)
	Intra-rater reliability	
ACSID	Percentage agreement=90.2-100.0 Cohen's κ=0.77-0.91 ICC (CI_{95})=0.84 (0.76-0.89)	Werner et al. (2018)

5x STS: Five Times Sit-to-Stand Test, 6min WT: 6-minute walk test, ACSID: Assessment of Compensatory Sit-to-Stand Maneuvers in People With Dementia, BBS: Berg Balance Scale, CI_{95}: 95 % confidence interval, E-ADL Test: Erlangen Test of Activities of Daily Living, ICC: intraclass correlation coefficient, MMSE: Mini-Mental-State Examination, n.s.: not significant, OR: odds ratio, POMA: Performance Oriented Mobility Assessment, SPPB: Short Physical Performance Battery, STS: Sit-to-Stand, TUG: Timed Up & Go Test

Inter-rater reliability

***Table E.** Data extraction main search - inter-rater reliability

Motor assessment	Variable	Relative inter-rater reliability	Absolute inter-rater reliability	Reference
		Balance		
FR	Distance [cm]	ICC (CI_{95})=0.79 (0.43-0.94)	N/A	Muir-Hunter et al. (2015)
GMWT	Time [s]	ICC=0.99	SEM=1.00 MDC_{95}=2.78 $MDC_{95\%}$=14.5 %	H.-S. Lee et al. (2017)
	Number of oversteps	ICC=0.99	SEM=0.76 MDC_{95}=2.12 $MDC_{95\%}$=17.1 %	H.-S. Lee et al. (2017)
BBS	Score	ICC=0.99	SEM=0.78 MDC_{95}=2.18 $MDC_{95\%}$=5.1 %	H.-S. Lee et al. (2017)
		ICC (CI_{95})=0.72 (0.31-0.91)	N/A	Muir-Hunter et al. (2015)
		Weighted κ=0.94 ICC=0.99	SEM=0.97 MDC_{95}=1.92 $MDC_{95\%}$=7.0 %	Telenius et al. (2015b)
		Mobility and gait		
TUG	Time [s]	ICC=0.99	SEM=0.63 MDC_{95}=1.75 $MDC_{95\%}$=7.9 %	H.-S. Lee et al. (2017)
		ICC (CI_{95})=0.98 (0.93-0.99)	N/A	Muir-Hunter et al. (2015)
6m WT	Walking speed [m/s]	ICC=0.97	SEM=0.03 MDC_{95}=0.06 $MDC_{95\%}$=15.2 %	Telenius et al. (2015b)

4m WT	Walking speed [m/s]	ICC=0.82	SEM=0.74 MDC_{95}=2.06 $MDC_{95\%}$=98.0 %	H.-S. Lee et al. (2017)
		Strength		
30s CST	Repetitions	ICC=1.00	SEM=0.00 MDC_{95}=0.00 $MDC_{95\%}$=0.0 %	Telenius et al. (2015b)
ACSID	Score	Percentage agreement=92.1-100.0 Cohen's κ=0.64-0.99 ICC (CI_{95})=0.85 (0.78-0.90)	N/A	Werner et al. (2018)
		Endurance		
6min WT	Distance [ft]	AM: ICC=0.99 PM: ICC=0.97	N/A	Tappen et al. (1997)
	Walking speed [ft/s]	AM: ICC=0.98 PM: ICC=0.96	N/A	Tappen et al. (1997)

4m WT: 4-metre walk test, 6m WT: 6-metre walk test, 6min WT: 6-minute walk test, 30s CST: 30-second chair stand test, ACSID: Assessment of Compensatory Sit-to-Stand Maneuvers in People With Dementia, AM: morning measures, BBS: Berg Balance Scale, CI_{95}: 95 % confidence interval, FR: Functional Reach Test, GMWT: Groningen Meander Walking Test, ICC: intraclass correlation coefficient, MDC_{95}: minimal detectable changes at 95 % confidence interval, $MDC_{95\%}$: percentage minimal detectable changes at 95 % confidence interval, N/A: not applicable, PM: afternoon measures, SEM: standard error of measurement, TUG: Timed Up & Go Test

Test-retest reliability

***Table F.** Data extraction main search - test-retest reliability

Motor assessment	Variable	Relative test-retest reliability	Absolute test-retest reliability	Reference
		Balance		
FICSIT-4	Score	ICC (CI_{95})=0.79 (0.67-0.87)	SEM (CI_{95})=0.55 (0.47-0.69) MDC_{95}=1.52 $MDC_{95\%}$=59.4 %	Blankevoort et al. (2013)
		ICC (CI_{95})=0.82 (0.65-0.91)	SEM (CI_{95})=0.59 (0.48-0.81) MDC_{95}=1.64 $MDC_{95\%}$=58.9 %	Blankevoort et al. (2013)[SG1]
		ICC (CI_{95})=0.80 (0.61-0.90)	SEM (CI_{95})=0.60 (0.48-0.82) MDC_{95}=1.66 $MDC_{95\%}$=71.1 %	Blankevoort et al. (2013)[SG2]
Modified Clinical Test of Sensory Interaction of Balance	Sway velocity [deg/s]	ICC=0.91	SEM=0.17 MDC_{95}=0.34 $MDC_{95\%}$=36.5 % CV=14.9 %	Suttanon et al. (2011)
Limits of Stability	Reaction time [s]	ICC=0.52	SEM=0.15 MDC_{95}=0.29 $MDC_{95\%}$=38.0 % CV=14.2 %	Suttanon et al. (2011)
	Movement velocity [deg/s]	ICC=0.48	SEM=0.46 MDC_{95}=0.91 $MDC_{95\%}$=38.9 % CV=14.7 %	Suttanon et al. (2011)

	Maximum excursion [%]	ICC=0.68	SEM=4.44 MDC_{95}=8.71 $MDC_{95\%}$=15.9 % CV=6.2 %	Suttanon et al. (2011)
	Directional control [%]	ICC=0.71	SEM=5.24 MDC_{95}=10.27 $MDC_{95\%}$=21.8 % CV=8.3 %	Suttanon et al. (2011)
Physiomat-Trail-Making Task	Score	r_s=0.89 ICC (CI_{95})=0.90 (0.85-0.95)	N/A	Wiloth et al. (2016)
Physiomat-Trail-Making Task simple	Sway path [mm/s]	r_s=0.59 ICC (CI_{95})=0.47 (0.27-0.63)	N/A	Wiloth et al. (2016)
	Time [s]	r_s=0.60 ICC (CI_{95})=0.55 (0.37-0.69)	N/A	Wiloth et al. (2016)
Physiomat-Trail-Making Task moderate	Sway path [mm/s]	r_s=0.78 ICC (CI_{95})=0.74 (0.61-0.82)	N/A	Wiloth et al. (2016)
	Time [s]	r_s=0.74 ICC (CI_{95})=0.79 (0.68-0.87)	N/A	Wiloth et al. (2016)
Physiomat-Trail-Making Task complex	Sway path [mm/s]	r_s=0.80 ICC (CI_{95})=0.82 (0.69-0.89)	N/A	Wiloth et al. (2016)
	Time [s]	r_s=0.87 ICC (CI_{95})=0.83 (0.72-0.91)	N/A	Wiloth et al. (2016)
Physiomat-Follow-the-Ball Task	Sway path [mm/s]	r_s=0.74 ICC (CI_{95})=0.84 (0.76-0.89)	N/A	Wiloth et al. (2016)
	Duration [s]	r_s=0.69 ICC (CI_{95})=0.79 (0.68-0.86)	N/A	Wiloth et al. (2016)
FR	Distance [cm]	ICC (CI_{95})=0.81 (0.52-0.94)	SEM=4.56 MDC_{95}=12.64 $MDC_{95\%}$=68.9 %	Muir-Hunter et al. (2015)

		ICC=0.84	SEM=1.61 MDC_{95}=3.15 $MDC_{95\%}$=15.4 % CV=5.7 %	Suttanon et al. (2011)
Hill Step Test	Number of steps (worst leg)	ICC=0.87	SEM=1.24 MCD_{95}=2.42 $MDC_{95\%}$=26.2 % CV=11.3 %	Suttanon et al. (2011)
Step Quick Turn	Time [s]	ICC=0.55	SEM=0.33 MDC_{95}=0.64 $MDC_{95\%}$=38.1 % CV=14.4 %	Suttanon et al. (2011)
	Sway [deg/s]	ICC=0.64	SEM=4.56 MDC_{95}=8.93 $MDC_{95\%}$=29.7 % CV=10.5 %	Suttanon et al. (2011)
Figure of Eight	Time [s]	ICC (CI_{95})=0.91 (0.85-0.95)	SEM (CI_{95})=6.26 (5.41-8.21) MDC_{95}=17.35 $MDC_{95\%}$=37.9 %	Blankevoort et al. (2013)
		ICC (CI_{95})=0.94 (0.86-0.97)	SEM (CI_{95})=6.24 (5.63-10.03) MDC_{95}=17.30 $MDC_{95\%}$=36.9 %	Blankevoort et al. (2013)[SG1]
		ICC (CI_{95})=0.85 (0.67-0.94)	SEM (CI_{95})=6.00 (4.01-7.58) MDC_{95}=16.63 $MDC_{95\%}$=37.4 %	Blankevoort et al. (2013)[SG2]
GMWT	Time [s]	ICC (CI_{95})=0.94 (0.90-0.97)	SEM (CI_{95})=1.93 (1.64-2.54) MDC_{95}=5.35 $MDC_{95\%}$=31.2 %	Bossers, van der Woude et al. (2014)

		ICC=0.96	n.r.	Bossers, van der Woude et al. (2014)[SG1]
		ICC=0.93	n.r.	Bossers, van der Woude et al. (2014)[SG2]
		ICC=0.99	SEM=1.36 MDC_{95}=3.78 $MDC_{95\%}$=19.6 %	H.-S. Lee et al. (2017)
	Number of oversteps	ICC (CI_{95})=0.63 (0.41-0.78)	SEM (CI_{95})=1.58 (1.31-2.03) MDC_{95}=4.38 $MDC_{95\%}$=225.7 %	Bossers, van der Woude et al. (2014)
		ICC=0.79	n.r.	Bossers, van der Woude et al. (2014)[SG1]
		ICC=0.57	n.r.	Bossers, van der Woude et al. (2014)[SG2]
		ICC=0.96	SEM=1.49 MDC_{95}=4.13 $MDC_{95\%}$=33.3 %	H.-S. Lee et al. (2017)
BBS	Score	ICC=0.99	SEM=1.36 MDC_{95}=3.78 $MDC_{95\%}$=10.2 %	H.-S. Lee et al. (2017)
		ICC (CI_{95})=0.95 (0.85-0.98)	SEM=6.01 MDC_{95}=16.66 $MDC_{95\%}$=38.6 %	Muir-Hunter et al. (2015)
			Mobility and gait	
TUG	Time [s]	ICC (CI_{95})=0.94 (0.92-0.97)	SEM (CI_{95})=2.12 (1.74-2.52) MDC_{95}=5.88 $MDC_{95\%}$=31.6 %	Blankevoort et al. (2013)
		ICC (CI_{95})=0.96 (0.92-0.98)	SEM (CI_{95})=1.43 (1.06-1.79) MDC_{95}=3.96 $MDC_{95\%}$=23.3 %	Blankevoort et al. (2013)[SG1]

		ICC (CI_{95})=0.94 (0.87-0.97)	SEM (CI_{95})=2.91 (2.10-3.61) MDC_{95}=8.07 $MDC_{95\%}$=39.6 %	Blankevoort et al. (2013)[SG2]
		ICC=0.99	SEM=1.27 MDC_{95}=3.52 $MDC_{95\%}$=15.8 %	H.-S. Lee et al. (2017)
		ICC=0.99	SEM=2.48 MDC_{90}=4.09 $MDC_{95\%}$=27.7 %	Ries et al. (2009)
		ICC=0.99	SEM=1.52 $MDC_{95\%}$=21.1 %	Ries et al. (2009)[SG1/2]
		ICC=99	SEM=3.03 $MDC_{95\%}$=30.0 %	Ries et al. (2009)[SG3]
		ICC (CI_{95})=0.72 (0.33-0.90)	SEM=1.24 MDC_{95}=3.44 $MDC_{95\%}$=20.3 %	Muir-Hunter et al. (2015)
		ICC=0.76	SEM=1.24 MDC_{95}=2.42 $MDC_{95\%}$=24.9 % CV=9.4 %	Suttanon et al. (2011)
		ICC=0.87	N/A	Thomas and Hageman (1999)
Cognitive TUG	Time [s]	ICC=0.51	SEM=2.39 MDC_{95}=4.69 $MDC_{95\%}$=36.2 % CV=14.1 %	Suttanon et al. (2011)
Manual TUG	Time [s]	ICC=0.70	SEM=1.45 MDC_{95}=2.83 $MDC_{95\%}$=26.7 % CV=10.1 %	Suttanon et al. (2011)

6m WT (comfortable pace)	Walking speed [m/s]	ICC (CI_{95})=0.86 (0.78-0.92)	SEM (CI_{95})=0.10 (0.08-0.12) MDC_{95}=0.27 $MDC_{95\%}$=36.5 %	Blankevoort et al. (2013)
		ICC (CI_{95})=0.83 (0.67-0.91)	SEM (CI_{95})=0.11 (0.09-0.11) MDC_{95}=0.29 $MDC_{95\%}$=41.5 %	Blankevoort et al. (2013)[SG1]
		ICC (CI_{95})=0.89 (0.78-0.95)	SEM (CI_{95})=0.09 (0.07-0.13) MDC_{95}=0.25 $MDC_{95\%}$=31.6 %	Blankevoort et al. (2013)[SG2]
	Time [s]	ICC=0.92	N/A	Thomas and Hageman (1999)
	Number of steps	ICC=0.80	N/A	Thomas and Hageman (1999)
6m WT (fast pace)	Time [s]	ICC=0.95	N/A	Thomas and Hageman (1999)
	Number of steps	ICC=0.90	N/A	Thomas and Hageman (1999)
4m WT	Walking speed [m/s]	ICC=0.85	SEM=0.64 MDC_{95}=1.78 $MDC_{95\%}$=84.3 %	H.-S. Lee et al. (2017)
Gait analysis (GAITRite)	Walking speed [m/s]	ICC (CI_{95})=0.95 (0.81-0.99)	N/A	McGough et al. (2013)
	Walking speed [cm/s]	ICC=0.98	SEM=5.72 MDC_{90}=9.44 $MDC_{95\%}$=27.4 %	Ries et al. (2009)
		ICC=0.97-0.98*	SEM=6.07 $MDC_{95\%}$=25.5 %	Ries et al. (2009)[SG1/2]
		ICC=0.97-0.98*	SEM=5.48 $MDC_{95\%}$=29.0 %	Ries et al. (2009)[SG3]

Walking speed [m/s]	ICC (CI_{95})=0.95 (0.88–0.98)	MCD_{95}=0.13 $MDC_{95\%}$=12.0 % CV=4.2 %	Wittwer et al. (2008)[3 walks]
	ICC (CI_{95})=0.96 (0.91–0.99)	MCD_{95}=0.11 $MDC_{95\%}$=10.2 % CV=3.8 %	Wittwer et al. (2008)[10 walks]
Step length [cm]	ICC (CI_{95})=0.97 (0.93-0.99)	MCD_{95}=5.27 $MDC_{95\%}$=8.9 % CV=3.1 %	Wittwer et al. (2008)[3 walks, r]
	ICC (CI_{95})=0.98 (0.96-0.99)	MCD_{95}=4.15 $MDC_{95\%}$=7.0 % CV=2.5 %	Wittwer et al. (2008)[10 walks, r]
Step width [cm]	ICC (CI_{95})=0.92 (0.82-0.97)	MCD_{95}=2.23 $MDC_{95\%}$=24.7 % CV=8.9 %	Wittwer et al. (2008)[3 walks, r]
	ICC (CI_{95})=0.95 (0.87-0.98)	MCD_{95}=1.83 $MDC_{95\%}$=20.0 % CV=7.0 %	Wittwer et al. (2008)[10 walks, r]
Stride length (cm)	ICC (CI_{95})=0.97 (0.87-0.99)	N/A	McGough et al. (2013)
	ICC (CI_{95})=0.97 (0.93-0.99)	MCD_{95}=10.24 $MDC_{95\%}$=8.5 % CV=3.0 %	Wittwer et al. (2008)[3 walks, r]
	ICC (CI_{95})=0.98 (0.96-0.99)	MCD_{95}=8.12 $MDC_{95\%}$=6.8 % CV=2.4 %	Wittwer et al. (2008)[10 walks, r]
Cadence [steps/ min]	ICC (CI_{95})=0.91 (0.62-0.98)	N/A	McGough et al. (2013)
	ICC (CI_{95})=0.88 (0.72–0.95)	MCD_{95}=8.13 $MDC_{95\%}$=7.5 % CV=2.7 %	Wittwer et al. (2008)[3 walks]

	ICC (CI_{95})=0.89 (0.74–0.95)	MDC_{95}=7.64 $MDC_{95\%}$=7.1 % CV=2.5 %	Wittwer et al. (2008)[10 walks]
Swing time [s]	ICC (CI_{95})=0.96 (0.81-0.99)	N/A	McGough et al. (2013)
	ICC (CI_{95})=0.90 (0.76–0.96)	MCD_{95}=0.03 $MDC_{95\%}$=7.1 % CV=2.7 %	Wittwer et al. (2008)[3 walks, r]
	ICC (CI_{95})=0.89 (0.75–0.96)	MCD_{95}=0.03 $MDC_{95\%}$=7.0 % CV=2.8 %	Wittwer et al. (2008)[10 walks, r]
Stance time [s]	ICC (CI_{95})=0.87 (0.70-0.95)	MCD_{95}=0.06 $MDC_{95\%}$=8.7 % CV=3.3 %	Wittwer et al. (2008)[3 walks, r]
	ICC (CI_{95})=0.88 (0.73-0.95)	MCD_{95}=0.06 $MDC_{95\%}$=8.6 % CV=2.9 %	Wittwer et al. (2008)[10 walks, r]
Toe in/out angle [deg]	ICC (CI_{95})=0.91 (0.78-0.96)	MCD_{95}=3.06 $MDC_{95\%}$=33.5 % CV=12.9 %	Wittwer et al. (2008)[3 walks, r]
	ICC (CI_{95})=0.93 (0.82-0.97)	MCD_{95}=2.58 $MDC_{95\%}$=28.2 % CV=10.8 %	Wittwer et al. (2008)[10 walks, r]
Walking speed variability [%]	ICC (CI_{95})=0.66 (0.26-0.87)	SEM=1.60 MDC_{95}=4.40 $MDC_{95\%}$=77.8 %	Wittwer et al. (2013)
Stride length variability [%]	ICC (CI_{95})=0.80 (0.52-0.93)	SEM=1.10 MDC_{95}=3.10 $MDC_{95\%}$=71.7 %	Wittwer et al. (2013)

	Stride width variability [%]	ICC (CI_{95})=0.83 (0.59-0.94)	SEM=3.00 MDC_{95}=8.30 $MDC_{95\%}$=47.0 %	Wittwer et al. (2013)
	Cadence variability [%]	ICC (CI_{95})=0.65 (0.25-0.86)	SEM=0.80 MDC_{95}=2.30 $MDC_{95\%}$=41.4 %	Wittwer et al. (2013)
Gait analysis (NeuroCom Balance Master)	Walking speed [cm/s]	ICC=0.50	SEM=7.58 MDC_{95}=14.86 $MDC_{95\%}$=48.3 % CV=20.6 %	Suttanon et al. (2011)
	Step length [cm]	ICC=0.75	SEM=4.59 MDC_{95}=9.00 $MDC_{95\%}$=35.6 % CV=13.9 %	Suttanon et al. (2011)
	Step width [cm]	ICC=0.89	SEM=1.26 MDC_{95}=2.48 $MDC_{95\%}$=22.0 % CV=14.7 %	Suttanon et al. (2011)
		Strength		
5x STS	Time [s]	ICC=0.80	SEM=1.39 MCD_{95}=2.73 $MDC_{95\%}$=29.9 % CV=10.5 %	Suttanon et al. (2011)
		ICC=0.94	N/A	Thomas and Hageman (1999)
STS on NeuroCom Balance Master	Rising index [% body weight]	ICC=0.95	SEM=1.25 MCD_{95}=2.44 $MDC_{95\%}$=21.8 % CV=7.7 %	Suttanon et al. (2011)

	COG sway velocity [deg/s]	ICC=0.02	SEM=1.20 MCD_{95}=2.35 $MDC_{95\%}$=80.2 % CV=39.2 %	Suttanon et al. (2011)
Modified 30s CST	Repetitions	ICC (CI_{95})=0.84 (0.73-0.90)	SEM (CI_{95})=1.26 (1.06-1.57) MDC_{95}=3.49 $MDC_{95\%}$=42.5 %	Blankevoort et al. (2013)
		ICC (CI_{95})=0.79 (0.60-0.90)	SEM (CI_{95})=1.52 (1.22-2.08) MDC_{95}=4.21 $MDC_{95\%}$=45.7 %	Blankevoort et al. (2013)[SG1]
		ICC (CI_{95})=0.88 (0.73-0.95)	SEM (CI_{95})=0.83 (0.65-1.04) MDC_{95}=2.30 $MDC_{95\%}$=33.2 %	Blankevoort et al. (2013)[SG2]
Handgrip dynamometer	Force [kgf]	ICC=0.98	N/A	Alencar et al. (2012)[SG0]
		ICC=0.97	N/A	Alencar et al. (2012)[SG1]
		ICC=0.96	N/A	Alencar et al. (2012)[SG2]
		ICC=0.42	N/A	Alencar et al. (2012)[SG3]
	Force [kg]	ICC (CI_{95})=0.90 (0.84-0.94)	SEM (CI_{95})=2.74 (2.05-2.98) MDC_{95}=7.59 $MDC_{95\%}$=36.8 %	Blankevoort et al. (2013)
		ICC (CI_{95})=0.86 (0.72-0.93)	SEM (CI_{95})=2.75 (1.85-3.15) MDC_{95}=7.62 $MDC_{95\%}$=36.5 %	Blankevoort et al. (2013)[SG1]
		ICC (CI_{95})=0.94 (0.87-0.97)	SEM (CI_{95})=2.57 (2.02-3.47) MDC_{95}=7.11 $MDC_{95\%}$=34.9 %	Blankevoort et al. (2013)[SG2]
	Force [kg]	Right: ICC=0.68 Left: ICC=0.70	N/A	Thomas and Hageman (1999)

Maximal isometric strength assessed with dynamometers (knee extension strength)	Torque [Nm]/normalized torque [Nm/kg]	ICC=0.97		Suzuki et al. (2009)
		ICC=0.98	N/A	Suzuki et al. (2009)[SG1/2]
		ICC=0.95	N/A	Suzuki et al. (2009)[SG3]
	Peak force [kgf]	Right: ICC=0.63 Left: ICC=0.56	N/A	Thomas and Hageman (1999)
Maximal isometric strength assessed with dynamometers (hip flexor strength)	Peak force [kgf]	Right: ICC=0.71 Left: ICC=0.62	N/A	Thomas and Hageman (1999)
Maximal isometric strength assessed with dynamometers (dorsiflexor muscles strength)	Peak force [kgf]	Right: ICC=0.63 Left: ICC=0.77	N/A	Thomas and Hageman (1999)
		Endurance		
6min WT	Distance [m]	ICC=0.99	SEM=20.28 MDC_{90}=33.47 $MDC_{95\%}$=23.9 %	Ries et al. (2009)
		ICC=0.98-0.99*	SEM=21.86 $MDC_{95\%}$=21.2 %	Ries et al. (2009)[SG1/2]
		ICC=0.98-0.99*	SEM=19.57 $MDC_{95\%}$=28.9 %	Ries et al. (2009)[SG3]
	Distance [ft]	Examiner 1, week 1: ICC=0.90 Examiner 1, week 2: ICC=0.80 Examiner 2, week 2: ICC=0.84	N/A	Tappen et al. (1997)
		AM: ICC=0.84 PM: ICC=0.76	N/A	Tappen et al. (1997)
	Walking speed [ft/s]	Examiner 1, week 1: ICC=0.89 Examiner 1, week 2: ICC=0.79 Examiner 2, week 2: ICC=0.84	N/A	Tappen et al. (1997)

		AM: ICC=0.84 PM: ICC=0.75	N/A	Tappen et al. (1997)
		Functional performance		
E-ADL Test	Score	r=0.73 (items: r=0.35-0.63)	N/A	Graessel et al. (2009)

4m WT: 4-metre walk test, 5x STS: Five Times Sit-to-Stand Test, 6m WT: 6-metre walk test, 6min WT: 6-minute walk test, 30s CST: 30-second chair stand test, AM: morning measures, BBS: Berg Balance Scale, CI_{95}: 95 % confidence interval, COG: centre of gravity, CV: coefficient of variation, E-ADL Test: Erlangen Test of Activities of Daily Living, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, FR: Functional Reach Test, GMWT: Groningen Meander Walking Test, ICC: intraclass correlation coefficient, kgf: kilogram-force, MDC_{90}: minimal detectable changes at 90 % confidence interval, MDC_{95}: minimal detectable changes at 95 % confidence interval, $MDC_{95\%}$: percentage minimal detectable changes at 95 % confidence interval, N/A: not applicable, n.r.: not reported, PM: afternoon measures, SEM: standard error of measurement, SG: subgroup, STS: Sit-to-Stand, TUG: Timed Up & Go Test

* range of ICC for several subgroups, no exact ICC reported

Additional file 9

Description, frequency of use, and effect sizes of motor assessments applied in previous randomised controlled trials without available information on psychometric properties

***Table G.** Description, frequency of use, and effect sizes of motor assessments applied in randomised controlled trials (no psychometric properties investigations available)

Motor Assessment	Description	Frequency of use	Time*group interaction effect size
	Balance		
Near-tandem test	*Task*: adopting near-tandem position with eyes closed (not exactly specified) *Measurement*: not specified	1 RCT (n=19) (Wesson et al., 2013)	-
Single leg stance/One Leg Standing Balance Test (Hawk, Hyland, Rupert, Colonvega, & Hall, 2006)	*Task*: standing on a single leg alternately for 60/30 seconds with both eyes open and closed *Measurement*: time [s]	2 RCT (n=63) (Burgener et al., 2008; Kampragkou et al., 2017)	Large [c]
One-leg balance test (Vellas et al., 1997)	*Task*: standing on one leg unsupported for five seconds (preferred leg) *Measurement*: recording if participant is able to maintain one-leg stance for five seconds [≥5s: normal, <5s: abnormal]	1 RCT (n=110) (Rolland et al., 2007)	-
Inertial sensors assessing postural sway (Moe-Nilssen, 1998)	*Task*: standing quietly with feet together for 30 seconds, while wearing an inertial sensor (DynaPort) *Measurement*: sway area [sq cm]	1 RCT (n=81) (Schwenk, Dutzi et al., 2014)	Small/ medium [r]
Posturography platforms assessing postural sway	*Task*: standing quietly on a posturography platform (QFP) for 51.2 seconds with eyes open *Measurement*: elliptical area covered by moving centre of gravity [mm²]	1 RCT (n=20) (Toulotte et al., 2003)	-
Wii Balance Board assessing postural sway (Clark et al., 2010)	*Task*: standing quietly on a Wii Balance Board for 15/30 seconds under four conditions: feet apart with eyes open and closed, feet together with eyes open and closed *Measurement*: centre or pressure velocity [cm/s], centre of pressure path length [cm/s]	1 RCT (n=20) (Yoon et al., 2013)	-
	Mobility and gait		
Get-Up and Go Test (Mathias et al., 1986)	*Task*: standing up from a chair, walking three metres, turning around, walking back to the chair, and sitting down, use of a walking aid is allowed *Measurement*: score from 1 to 5 [1=no instability to 5=very abnormal]	1 RCT (n=110) (Rolland et al., 2007)	-

10-meter walk test (Guralnik, Seeman et al., 1994)	*Task*: walking ten metres with comfortable pace, use of walking aid is allowed *Measurement*: walking speed [m/s], time [s]	2 RCT (n=209) (Sobol et al., 2016; Toulotte et al., 2003)	-
10-meter walk test with dual Task (Guralnik, Seeman et al., 1994)	*Task*: walking ten metres with comfortable pace while naming month backwards/counting backwards from 50, use of walking aid is allowed *Measurement*: walking speed [m/s]	1 RCT (n=189) (Sobol et al., 2016)	-
8-foot walk test (Guralnik, Seeman et al., 1994)	*Task*: walking eight foot with comfortable/fast pace, use of walking aid is allowed *Measurement*: walking speed [m/s], time [s]	3 RCT (n=78) (Dawson et al., 2019; Hauer et al., 2017; Steinberg et al., 2009)	Medium [r]
Gait analysis with dual task performance (Kressig & Beauchet, 2006)	*Task*: walking with comfortable pace over an electronic walkway (GAITRite) while counting forward by twos/backward by threes *Measurement*: dual-task cost in motor performance [%] for walking speed, cadence, stride length, stride time, single support, dual-task cost in combined performance [%]	1 RCT (n=49) (Schwenk et al., 2010)	-
Southampton Assessment of Mobility (Pomeroy, 1990)	*Task*: different tasks of mobility (sit-to-stand, standing balance, gait, stand-to-sit) *Measurement*: score [0=immobile, 18=able to walk four steps]	1 RCT (n=78) (Pomeroy et al., 1999)	Small [c]
Hierarchical Assessment of Balance and Mobility (Rockwood, Rockwood, Andrew, & Mitnitski, 2008)	*Task*: getting up from bedside and walking, which is rated in three areas: in-bed mobility, transfers, and walking *Measurement*: score [0-65]	1 RCT (n=123) (Schwenk, Dutzi et al., 2014)	Small [r]
Acute Care Index of Function (Roach & van Dillen, 1988)	*Task*: 20 items that are divided into four subscales - mental status, bed mobility, transfers, and mobility *Measurement*: score [0-1]	1 RCT (n=82) (Roach et al., 2011)	-
	Strength		
Body-fixed-sensor-based Sit-to-Stand analysis using DynaPort (Schwenk et al., 2012)	*Task*: performing five repetitions of the sit-to-stand task without upper extremity assistance *Measurement*: total duration [s], duration of hip extension/flexion [s], maximum angular velocity during hip extension/flexion [deg/s] of stand-to-sit and sit-to-stand movements and trunk flexion range [deg], trunk flexion duration [s], maximum trunk flexion angular velocity [deg/s], sit-to-stand movement duration	2 RCT (n=106) (Hauer et al., 2017; Werner et al., 2017)	No to large [c/r]
10-s chair-stand test (Bohannon, 1995)	*Task*: performing as many repetitions of sit-to-stand task as possible in 10 seconds without upper extremity assistance *Measurement*: number of repetitions	1 RCT (n=40) (Vreugdenhil et al., 2012)	Large [c]

One-repetition maximum assessed with fitness machines	*Task*: one-repetition maximum as achieved in the leg-press training machine for maximum dynamic concentric muscle strength in hip and knee extensors and in the abductor training machine for maximum strength in hip abductors, respectively *Measurement*: one-repetition maximum [kg]	2 RCT (n=232) (Hauer et al., 2012; Schwenk, Dutzi et al., 2014)	Medium/ large [r]
Stair-climbing performance (Reuben & Siu, 1990)	*Task*: climbing a flight with thirteen stairs *Measurement*: time [s]	1 RCT (n=107) (Hauer et al., 2012)	-
Physical therapy assessment	*Task*: measure muscle strength in upper and lower extremities (not exactly specified) *Measurement*: score [not specified]	1 RCT (n=11) (Francese et al., 1997)	-
	Endurance		
2-min walk test (Stewart, Burns, Dunn, & Roberts, 2016)	*Task*: walking for two minutes with comfortable pace, use of usual walking aids is allowed *Measurement*: distance [m]	2 RCT (n=152) (Cott et al., 2002; Pomeroy et al., 1999)	-
400-m walk test (Rolland et al., 2004)	*Task*: walking 400 m with fast pace *Measurement*: time [s]	1 RCT (n=189) (Sobol et al., 2016)	-
3-speed walking test	*Task*: walking on a treadmill with 80 %, 100 %, and 120 % of self-selected pace for five minutes *Measurement*: cost of walking [J/kg/m], VO_2 [ml/kg/min], heart rate [bpm]	1 RCT (n=34) (Pedrinolla et al., 2018)	Medium to large [c]
6-min Astrand Cycle Ergometer test (Astrand & Ryhming, 1954)	*Task*: submaximal 6-minute cycle test *Measurement*: estimated maximum oxygen uptake [ml/kg/min]	1 RCT (n=189) (Sobol et al., 2016)	-
Pedal Power	*Task*: pedalling in seven steps from 10 to 70 W (not exactly specified) *Measurement*: exercise time [s], pedal rotations [number], total load [W/s]	1 RCT (n=31) (M.-J. Kim et al., 2016)	-
	Flexibility		
Chair sit and reach (Jones, Rikli, Max, & Noffal, 1998)	*Task*: stretching one leg keeping heel on the floor and trying to touch the toes with the fingers while sitting on a chair *Measurement*: distance between the fingers and toes [cm]	1 RCT (n=20) (Toulotte et al., 2003)	-
	Functional performance		
Physical Performance Test (Reuben & Siu, 1990)	*Task*: seven items (writing a sentence, transferring five beans from a bowl to a cup with a teaspoon, lifting a book onto a shelf, putting on a coat, picking up a coin from the floor, walking 50 feet, and turning 360° while standing in one place), which are scored on a 4-point Likert scale *Measurement*: score [0-28]	1 RCT (n=105) (Bossers et al., 2016)	-

Senior Fitness Test (Rikli & Jones, 2006)	*Task*: battery of tests including: (1) muscle dynamic strength endurance of legs (30-s chair stand test) and upper body (arm curl test), (2) flexibility of lower (chair sit-and-reach test) and upper body (back scratch test), (3) speed, agility and balance while moving (8-foot up-and-go test), (4) aerobic endurance (2-minute step test) *Measurement*: (1) number of repetitions, (2) distance between toes and fingers / fingers [cm], (3) time [s], (4) number of repetitions	1 RCT (n=16) (Santana-Sosa et al., 2008)	Large [c/r]
Jebsen Total Time (Jebsen, Taylor, Trieschmann, Trotter, & Howard, 1969)	*Task*: performing a range of seven hand functions required for activities of daily living (writing, turning over 3 by 5 inch cards, picking up small common objects, simulated feeding, stacking checkers; picking up large objects, and picking up large heavy objects), while total time is evaluated *Measurement*: total time [s]	1 RCT (n=27) (Steinberg et al., 2009)	-
Physiological Profile Assessment (Lord, Menz, & Tiedemann, 2003)	*Task*: five items evaluating visual contrast sensitivity, knee joint proprioception, quadriceps strength, simple reaction time, and postural sway while standing on a foam rubber mat with eyes open *Measurement*: fall risk score [z-score]	2 RCT (n=60) (Suttanon et al., 2013; Wesson et al., 2013)	-

n: number of analysed participants, RCT: randomised controlled trial/s

[c] calculated effect size, [r] effect size provided of randomised controlled trial

Multimedia Appendix manuscript III

Multimedia Appendix 1

Sample session

***Table H.** Arrival of a sample session of the multimodal exercise program

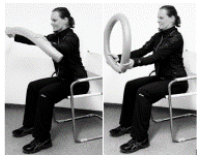
Imagination	Motor tasks	Cognitive tasks
Destination: ask the participants for destination of this session.	None.	If participants do not know destination, instructor gives explanations and descriptions about destination to create an imagination.
Pack your bag: take clothes out of the wardrobe.	 Move your arms and upper body with straightened back to the right, middle, and left side and then down to the floor, respectively (in order to put the clothes into the bag).	Instructor asks where in the wardrobe trousers (in middle compartment), T-shirts, and pullovers (in left overhead compartment), swimsuit or swim trunks (in right overhead compartment), etc. are.
Central Station: walk to the central station (meet the other tourists).	 Alternately lift your legs and swing your arms to simulate walking.	None.
Greet your fellow passengers.	 Upper body rotates to the left/right side and shake hands with fellow passenger.	None.

Take your ticket out of your pocket (trousers or shirt) or hand bag and show it to the conductor.	 Bend down to touch the chair leg (handbag) (alternatively touch the addressed pocket of your clothes), then stretch 1 arm with straightened back to the front and hold it.	Participants should remember where they have put the ticket after the last journey.
The group reflects on packing and thinks about if everything important is included.	 Resting and recovering.	Instructor asks participants what they packed in their bags a few minutes ago.
Get out of the train and walk to the hotel for check-in.	 Alternately lift your legs and swing your arms to simulate walking.	None.

***Table I.** Main part of a sample session of the multimodal exercise program (destination is a cruise in the Mediterranean Sea)

Imagination	Starting position	Motor tasks	Cognitive task	Time
1st Holiday Event – Visiting Pisa				
Leave the cruise ship, say goodbye to the captain.	 Sit up straight.	 Alternately Lift your legs and wave to the captain and salute him with a tip on your forehead.	Instructor starts a conversation while simulating walking asking if participants have been in Italy/Pisa. If yes: "Can you tell us about your adventures?" If no: "What's typical about Italy? Do you know some stereotypes?"	1 min
Climb up the tower of Pisa.	 Stand behind the chair (hold on to the back of the chair).	 Alternately Lift your legs to your upper body while slightly leaning backwards/to the sides (due to the angle of the tower).	Participants count the stairs forward beginning with 1 (up to 15).	1 min
Enjoy the amazing view.	 Stand behind the chair (hold on to the back of the chair).	 Turn around 360° (1 turn left, 1 turn right) (Keep holding on to the back of the chair if there is a risk of falling!)	None.	1x in each direction, 2 min
Climb down the tower of Pisa.		See above "climbing up".	Participants count the stairs backwards beginning with 15.	1 min

Return to the cruise ship and greet the captain.			None.	1 min
2nd Holiday Event: Aqua Fitness in the deck pool of the cruise ship				
Put on your swimsuit or swim trunks.	 Sit up straight.	Lift your right leg and move your hands – starting at your upper leg – over your knee along your lower leg to your ankle and then up again until you reach your hip (simulating dressing up). Repeat the task on your other leg.	None.	2x each leg, 0.5 min
Tasks with your pool noodle.	 Sit up straight, pool-noodle in both hands.		None.	10 min, 3 sets each task
		Hold pool-noodle with both hands (hands close to the ends). Stretch your arms vertically above your head. Put the pool-noodle behind your head – try to sit straight (it is allowed to bend the pool-noodle a little bit!).		3x
		 Hold pool-noodle with both hands (hands nearly at the ends). Stretch your arms vertically above your head. Lean your upper body to the right then to the left.		2x each side

		 Horizontally hold the pool-noodle (hands close to the ends). Put both ends together to form a circle. Maximally extend your arms.		approx. 20 sec
		 Bend the pool noodle in front of you (each end in 1 hand). Climb back and forth over the pool noodle (use it like a skipping rope).		3x back & forth
		 Vertically take the pool noodle between your legs, hold it in both hands with arms fully extended in front and roll it in your hands.		approx. 15 sec
The Pool Bar invites your tourist group for a drink (5min break for drinking)				5 min
3rd Holiday Event: Arrive in Cairo				
Leave the cruise ship, say goodbye to the captain and walk toward the camel station.		(See above “leave the cruise ship”).	None.	1 min

Get up on the (sitting) camel. Due to the superstition of the camel drivers, you have to walk around the camel 3 times before sitting down.	 Stand behind the chair (hold on to the back of the chair).	 Walk 3 times around your chair, change direction each time you finished a round.	None.	4 min
Ride the camel to the pyramids.	 Sit up straight.	 Take the reins → Arms are held in horizontal position while doing the other tasks. The camel moves slowly (wavers) → Hip tilts to the left/right/front/back while sitting on the chair. The camel gallops (hops) → rising a bit from the seat doing “ups” and “downs”.	Instructor tells the story about riding the camel above different barriers. At least 2 times participants have to ride slowly and gallop. Participants have to reply on the story by their movements.	2x 1 min activity 0.5min pause, 1.5 min
Arrive at the pyramids and climb off the camel (same ritual like getting up on the camel).		(See above “Get up on the [sitting] camel”).	None.	4 min
Meet local Bedouins at the pyramids, participate in a traditional dance → repetitive dance choreography.	 Sit up straight.	Common time (4 beats per cycle).	None.	10 min

		 1st cycle: (1) put your right hand on the left shoulder and (2) the left hand on the right shoulder (→ arms crossed), (3) bow your upper body toward your right neighbor, then (4) sit straight again.		
		2nd cycle: (5) bow your upper body toward your right neighbor, then (6) sit straight again, (7 & 8) clap your hands twice.		
		 3rd cycle: (1) clap your left hand on your upper right leg, then (2) your right hand on your upper left leg, (3 & 4) tamp twice with your right foot.		
		 4th cycle: (5 & 6) stamp twice with your left foot, (7 & 8) clap your hands twice.		
Get up on the (sitting) camel.		(See above “Get up on the [sitting] camel”).	Participants should remember “The ritual”.	4min

Ride back to the cruise ship.		(See above “Ride the camel to the pyramids”).	(See above “telling a Story”).	2x 1 min activity 0.5 min pause, 1.5 min
Climb off the camel.		(See above “Get up on the [sitting] camel”).	Together, instructor and participants verbally reconstruct the “ritual”.	4 min
Return to the cruise ship and greet the captain		Return to the cruise ship and greet the captain.	None.	1 min

***Table J.** Departure of a sample session of the multimodal exercise program

Imagination	**Motor tasks**	**Cognitive task**
Pack your bag.	None.	Participants have to remember the clothes and other things they have put in the bag on the arrival. Instructor encourages them by asking explicit questions.
Walk to the central station.	(See “Arrival”).	None.
Take your ticket out of your pocket (trousers or shirt) or hand bag and show it to the conductor.	(See “Arrival”).	Participants have to remember where they have left the ticket after they have showed it to the conductor.
Remembering.	 Relax and shake your arms and legs.	Participants have to remember what the destination of the journey today was and what experiences they had. Instructor encourages group to talk about the training lesson and if needed give hints.
Say goodbye to fellow passengers.	 Wave your hands at other participants.	None.
Announcing the next destination of the upcoming training lesson.	None.	Instructor says goodbye to participants and give a brief outlook to the next training session.

Multimedia Appendix 2

Description of the assessments

Multimedia Appendix 2 gives a detailed description of performing each assessment. These are important information to attain a standardized testing procedure and ensure comparability.

FICSIT-4 Scale (static balance; Rossiter-Fornoff et al., 1995)

Frailty and Injuries: Cooperative Studies of Intervention Techniques 4 scale (FICSIT-4) (Rossiter-Fornoff et al., 1995) determines static balance. Participants are asked to take different standing positions for 10 seconds each in the following order: Romberg, semi tandem, tandem, and single leg. First, the investigators demonstrate the positions and - if necessary - assist participants to take this position. Time is measured from the moment participants take up the position without help up to 10 seconds or in the following cases: foot position is changed or help is required to avoid a fall. Prerequisite for performing the next task of FICSIT-4 scale is the successful performance (10 seconds in the position) of the previous task. There is no test run before recording. The FICSIT-4 scale rates performance with 0 to 5 points according to number and time of finished positions (Rossiter-Fornoff et al., 1995).

Timed Up and Go test (mobility; Podsiadlo & Richardson, 1991)

Timed Up and Go test (Podsiadlo & Richardson, 1991) assess mobility. For timed Up and Go test participants are asked to rise from a chair, walk 3 meters, turn around, go back, and sit down on the chair. Time recording starts with “Go” and stops when participants sit on the chair again. The chair has a sitting height of 46 centimeters and armrests. The distance of 3 meters is marked with a cone. Using a walking aid is allowed and is placed next to participants. Investigators demonstrate timed Up and Go test once and participants do one test run. During assessment, 3 instructions are allowed where needed: 1. “Go to the cone”, 2. “Turn around”, 3. “Sit down”. 2 valid trials are recorded.

6-meter walk test (mobility; Graham et al., 2008)

The 6-meter walk test (Graham et al., 2008) assesses mobility and aims to capture normal gait speed. To reduce bias caused by the testing situation, participants are not explicitly informed about time keeping. A straight and flat distance of 6 meters is

marked. During time keeping, investigators try to avoid conversation. All walking aids used in everyday life are applied. The 6-meter walk test is repeated 2 or 3 times if necessary.

GAITRite (gait parameters)

Temporal and spatial gait parameters will be analyzed using the electronic gait analysis system GAITRite (CIR Systems Inc, Franklin, NJ) with an active length of 4.88 meters, a spatial resolution of 1.27 centimeters, and a scan rate of 120 hertz. Gait parameters are recorded for 3 different conditions: walking with normal speed; walking with normal speed and the task of counting backwards from 50; walking with normal speed and the task of naming animals. All conditions will be repeated up to 5 times walking in the same direction to generate 3 valid trials. All walking aids used in everyday life are applied. To eliminate acceleration and deceleration during recording, participants start walking 2 meters in front of the GAITRite system and end 2 meters after (Kressig & Beauchet, 2006). Rests between trials are allowed when necessary.

Modified 30-second chair-stand test (strength of lower limbs; Blankevoort et al., 2013; Jones et al., 1999)

The modified 30-second chair-stand test (Blankevoort et al., 2013; Jones et al., 1999) determines strength of lower limbs. Participants are asked to stand up and sit down as often as possible during 30 seconds. Repetitions are counted loudly. Moreover, the time to perform 5 repetitions is taken during the modified 30-second chair-stand test. In this modified version participants are allowed to use their arms (Blankevoort et al., 2013; Jones et al., 1999). The chair is the same as in timed Up and Go test (sitting height of 46 centimeters, with armrests). Investigators demonstrate the task and participants complete 1 test run. Valid performances, defined as hip angle during standing of about 180° and during sitting of about 90°, are counted after the command “Go” with simultaneous timing up to 30 seconds. If 30 seconds end while standing, a semi repetition is counted. After a rest, fit participants complete a second trial without using arms with the same recording procedure as for the modified 30-second chair-stand test (including time for 5 repetitions).

Short physical performance battery (function of lower limbs; Guralnik, Simonsick et al., 1994)

The short physical performance battery (Guralnik, Simonsick et al., 1994) evaluates function of lower limbs. It consists of standing balance (Romberg, semi tandem, tandem), gait speed, and 5 times sit to stand without using arms (Guralnik, Simonsick et al., 1994). All measures are described above.

Erlangen Test of Activities of Daily Living (E-ADL-Test) (Activities of Daily Living; Graessel et al., 2009)

Erlangen Test of Activities of Daily Living (E-ADL-Test) (Graessel et al., 2009) determines ADL. It consists of 5 items: pouring a drink, cutting a piece of bread, opening a small cupboard, washing hands, and tying a bow which will be performed during testing. A detailed description of each item is given by Graessel et al. (2009).

7-item physical performance test (Activities of Daily Living; Reuben & Siu, 1990)

The 7-item physical performance test (Reuben & Siu, 1990) assess ADLs and includes the following tasks: writing a sentence, simulated eating, turning 360 degrees, putting on and removing a jacket, lifting a book and putting it on a shelf, picking up a penny from the floor, and a 50-foot walk test. The 7-item physical performance test will be performed according to the test protocol given by Reuben and Siu (1990). Due to time restrictions and to reduce physical stress, the 50-foot walk test will not be performed in this high-aged sample and the gait speed of the 6-meter walk test will be used instead.

Cognitive Assessments

All cognitive assessments will be performed and rated according to available test protocols.

Body mass and height

Body mass and height will be measured using a Seca 813 Robusta scale and Seca 213 stadiometer (Seca, Hamburg, Germany) with an accuracy of 0.1 kilogram and 0.1 centimeter, respectively. Participants will wear normal clothes and shoes during all measurements and the shoe type will be documented.

Supplementary material manuscript V

Supplementary Table 1

***Table K.** Sample characteristics of participants at baseline (intention-to-treat analysis)

	Total sample [n=304]	**Intervention group** [n=194]	**Control group** [n=110]	**Group differences** [t(df)/z/Chi²(df), p]
Age, years [M (SD), range]	86 (6), 66-102	86 (6), 67-102	87 (6), 66-98	t(302)=1.135, p=0.257
Sex, female	86 %	85 %	89 %	Chi²(1)=1.223, p=0.269
Diagnosis of dementia				Chi²(2)=3.693, p=0.158
- yes	66 %	67 %	65 %	
- no	18 %	20 %	15 %	
- unknown	16 %	13 %	21 %	
Type of dementia				Chi²=9.005, p=0.050
- Alzheimer's disease	17 %	19 %	14 %	
- Vascular dementia	15 %	18 %	10 %	
- Mixed dementia	3 %	2 %	4 %	
- other	1 %	2 %	0 %	
- unknown	30 %	26 %	37 %	
- no/unknown diagnosis	34 %	34 %	36 %	
MMSE [M (SD), range]	17 (4), 10-24	17 (4), 10-24	17 (4), 10-24	t(250.853)=0.389, p=0.698
Use of walking aid				Chi²(2)=4.104, p=0.128
- walker	71 %	69 %	75 %	
- waking stick/s	6 %	4 %	8 %	
- no walking aid	21 %	24 %	16 %	
	unknown in 2 %	unknown in 3 %	unknown in 1 %	
Depression				Chi²(2)=2.461, p=0.292
- yes	26 %	28 %	22 %	
- no	52 %	53 %	52 %	
- unknown	22 %	20 %	26 %	
CIRS [M (SD), range]				
- Morbidity Index	9 (5), 1-26	9 (4), 1-20	10 (6), 2-26	t(176)=0.469, p=0.640
- Severity Index	1.6 (0.4), 1-3	1.6 (0.4), 1-3	1.6 (0.5), 1-3	z=-0.273, p=0.785
	not available for 41 %	not available for 37 %	not available for 50 %	
Number of medications [M (SD), range]	7 (4), 0-27	7 (4), 0-27	6 (4), 0-20	**t(232)=-2.686, p=0.008**
	unknown in 23 %	unknown in 22 %	unknown in 25 %	

Antidementives				Chi²(2)=2.742, p=0.254
- yes	22 %	25 %	17 %	
- no	43 %	41 %	47 %	
- unknown	35 %	34 %	36 %	
Antidepressants				**Chi²(2)=10.723, p=0.005**
- yes	25 %	31 %	15 %	
- no	36 %	32 %	42 %	
- unknown	39 %	37 %	44 %	
Height, cm [M (SD), range]	156.7 (8.1), 139.0-186.0 unknown in 6 %	156.6 (8.1), 139.0-186.0 unknown in 7 %	156.9 (8.3), 140.5-185.0 unknown in 5 %	t(284)=0.342, p=0.733
Weight, kg [M (SD), range]	68.9 (13.1), 41.3-125.0 unknown in 11 %	69.9 (13.5), 46.1-125.0 unknown in 12 %	67.2 (12.3), 41.3-99.4 unknown in 8 %	t(270)=-1.668, p=0.096
BMI, kg/m² [M (SD), range]	28.0 (4.7), 17.6-48.5 unknown in 11 %	28.5 (4.7), 18.1-48.5 unknown in 12 %	27.2 (4.8), 17.6-38.0 unknown in 9 %	**t(268)=-2.307, p=0.022**

BMI: Body Mass Index, CIRS: Cumulative Illness Rating Scale, df: degree of freedom, M: mean, MMSE: Mini-Mental State Examination, n: number, SD: standard deviation

Statistically significant results appear bold

Supplementary Table 2

***Table L.** Effects of the multimodal exercise program on spatiotemporal gait parameters and dual task costs (intention-to-treat analysis)

		Baseline [M (SE)]	**Group differences at baseline** [t(df), p]	**Post** [M (SE)]	**Difference post – baseline** [M (SE), [CI_{95}]]	**Within group time effects** [t(df), p]	**Time*group effects** $F(df_{numerator}, df_{denominator})$, p	Effect size η_p^2
Single task (IG: n=194, CG: n=110)								
Walking speed, m/sec	IG	0.62 (0.01)	t(136275)=-0.672, p=0.502	0.61 (0.01)	-0.01 (0.01), [-0.03, 0.02]	t(121)=0.573, p=0.568	F(1,302)=0.001 to 0.395, p=0.530 to 0.977	0.000 to 0.001
	CG	0.60 (0.02)		0.60 (0.02)	0.00 (0.02), [-0.03, 0.03]	t(238)=0.095, p=0.925		
Stride length, cm	IG	77.7 (1.4)	t(2596)=0.553, p=0.580	77.4 (1.6)	-0.3 (1.4), [-3.0, 2.5]	t(27)=0.190, p=0.851	F(1,302)=0.001 to 2.814, p=0.094 to 0.977 [a, b]	0.000 to 0.009
	CG	78.9 (1.7)		77.6 (1.9)	-1.3 (1.7), [-4.7, 2.0]	t(72)=0.789, p=0.433		
Stride time, sec	IG	1.3 (0.0)	**t(1200)=1.975, p=0.049**	1.3 (0.0)	0.0 (0.0), [0.0, 0.1]	t(46)=-1.224, p=0.227*	F(1,302)=0.325 to 4.944, p=0.027 to 0.569*	0.001 to 0.016
	CG	1.4 (0.0)		1.4 (0.0)	0.0 (0.0), [-0.1, 0.0]	t(230)=0.698, p=0.486		
Double support, % of stride time	IG	40.0 (0.6)	t(3372)=0.327, p=0.743	40.7 (0.6)	0.7 (0.6), [-0.4, 1.8]	t(59)=-1.318, p=0.192*	F(1,302)=0.002 to 0.265, p=0.607 to 0.961 [b]	0.000 to 0.001
	CG	40.3 (0.7)		41.2 (0.8)	0.9 (0.7), [-0.4, 2.2]	t(106)=-1.314, p=0.192*		
Stance phase, % of stride time	IG	69.8 (0.3)	t(1915)=0.305, p=0.760	70.2 (0.3)	0.4 (0.3), [-0.1, 1.0]	t(83)=-1.612, p=0.111*	F(1,302)=0.000 to 0.149, p=0.700 to 0.988	0.000
	CG	70.0 (0.4)		70.4 (0.5)	0.5 (0.4), [-0.3, 1.3]	t(48)=-1.143, p=0.259*		

Dual task, counting backwards (IG: n=194, KG: n=110)								
Walking speed, m/sec	IG	0.49 (0.01)	t(792)=-1.75, p=0.079*	0.49 (0.01)	0.00 (0.02), [-0.03, 0.03]	t(24)=0.035, p=0.972	F(1,302)=0.004 to 1.043, p=0.308 to 0.948	0.000 to 0.003
	CG	0.46 (0.02)		0.47 (0.02)	0.01 (0.02), [-0.02, 0.04]	t(185)=-0.540, p=0.590		
Stride length, cm	IG	71.2 (1.4)	t(1093)=-0.629, p=0.529	73.1 (1.6)	1.9 (1.6), [-1.4, 5.2]	t(20)=-1.194, p=0.246*	F(1,302)=0.063 to 1.212, p=0.272 to 0.802 [b]	0.000 to 0.004
	CG	69.8 (1.8)		72.7 (2.1)	2.9 (2.2), [-1.5, 7.4]	t(25)=-1.328, p=0.196*		
Stride time, sec	IG	1.5 (0.0)	t(108)=1.764, p=0.081*	1.6 (0.0)	0.0 (0.0), [0.0, 0.1]	t(30)=-0.836, p=0.410*	F(1,302)=0.000 to 4.518, p=0.034 to 0.996* [, a, b]	0.000 to 0.015
	CG	1.6 (0.0)		1.7 (0.1)	0.1 (0.1), [0.0, 0.2]	t(26)=-1.118, p=0.274*		
Double support, % of stride time	IG	43.8 (0.7)	t(212)=1.383, p=0.168	43.6 (0.8)	-0.2 (0.9), [-2.0, 1.6]	t(19)=0.237, p=0.815	F(1,302)=0.002 to 1.099, p=0.295 to 0.969	0.000 to 0.004
	CG	45.4 (0.9)		45.1 (1.0)	-0.3 (1.0), [-2.2, 1.7]	t(32)=0.275, p=0.785		
Stance phase, % of stride time	IG	71.7 (0.4)	t(73)=1.482, p=0.143*	71.8 (0.4)	0.1 (0.4), [-0.8, 1.0]	t(24)=-0.210, p=0.835	F(1,302)=0.035 to 2.770, p=0.097 to 0.852	0.000 to 0.009
	CG	72.7 (0.5)		72.4 (0.5)	-0.3 (0.6), [-1.5, 0.8]	t(29)=0.566, p=0.575*		
Dual-task costs, counting backwards (IG: n=194, KG: n=110)								
Walking speed, %	IG	-17.1 (2.0)	t(205)=1.649, p=0.101*	-15.4 (2.9)	1.7 (3.3), [-4.9, 8.4]	t(18)=-0.534, p=0.600	F(1,302)=0.032 to 2.378, p=0.124 to 0.859 [b]	0.000 to 0.008
	CG	-22.2 (2.0)		-17.8 (3.7)	4.4 (4.1), [-3.7, 12.5]	t(85)=-1.067, p=0.289		
Stride length, %	IG	-6.9 (1.5)	**t(93)=-2.192, p=0.031**	-3.5 (2.5)	3.4 (2.8), [-2.5, 9.4]	t(13)=-1.212, p=0.246*	F(1,302)=1.125 to 6.013, p=0.015 to 0.290* [, b]	0.004 to 0.015
	CG	-11.5 (1.4)		-3.6 (3.1)	7.9 (3.4), [1.1, 14.8]	**t(22)=-2.346, p=0.028**		
Stride time, %	IG	17.4 (2.1)	t(92)=0.513, p=0.609	18.6 (2.4)	1.2 (2.8), [-4.4, 6.9]	t(31)=-0.444, p=0.660	F(1,302)=0.000 to 4.709, p=0.031 to 0.998*	0.000 to 0.015
	CG	18.9 (2.2)		24.3 (3.5)	5.3 (3.8), [-2.2, 12.9]	t(27)=-1.414, p=0.169*		

Double support, %	IG	10.4 (1.2)	t(106)=1.572, p=0.119*	8.4 (1.8)	-2.0 (2.0), [-6.1, 2.1]	t(19)=1.004, p=0.328*	F(1,302)=0.000 to 0.715, p=0.398 to 0.998 [a, b]	0.000 to 0.002
	CG	13.4 (1.6)		11.0 (2.3)	-2.4 (2.5), [-7.4, 2.6]	t(32)=0.949, p=0.350*		
Stance phase, %	IG	2.8 (0.4)	t(39)=1.679, p=0.101*	2.4 (0.6)	-0.5 (0.7), [-1.8, 0.9]	t(21)=0.716, p=0.482*	F(1,302)=0.001 to 3.165, p=0.076 to 0.976 [b]	0.000 to 0.010
	CG	4.0 (0.5)		2.9 (0.7)	-1.1 (0.9), [-3.0, 0.7]	t(28)=1.222, p=0.232*		
Dual task, naming animals (IG: n=194, KG: n=110)								
Walking speed, m/sec	IG	0.41 (0.01)	t(103)=-0.706, p=0.482	0.42 (0.01)	0.01 (0.01), [-0.02, 0.03]	t(75)=-0.554, p=0.581	F(1,302)=0.002 to 0.974, p=0.324 to 0.967	0.000 to 0.003
	CG	0.40 (0.01)		0.41 (0.01)	0.01 (0.01), [-0.02, 0.04]	t(98)=-0.846, p=0.400		
Stride length, cm	IG	64.9 (1.4)	t(140)=0.394, p=0.694	67.7 (1.3)	2.8 (1.4), [0.0, 5.5]	t(25)=-2.009, p=0.056*	F(1,302)=0.001 to 3.912, p=0.049 to 0.971*	0.000 to 0.013
	CG	65.8 (1.7)		66.4 (1.6)	0.6 (1.9), [-3.2, 4.5]	t(31)=-0.316, p=0.754		
Stride time, sec	IG	1.7 (0.0)	t(134)=1.788, p=0.076*	1.7 (0.0)	0.0 (0.0), [0.0, 0.1]	t(28)=-0.971, p=0.340*	F(1,302)=1.921 to 10.040, p=0.002 to 0.167* [, a, b]	0.006 to 0.032
	CG	1.8 (0.0)		1.7 (0.1)	-0.1 (0.1), [-0.2, 0.1]	t(21)=1.003, p=0.327*		
Double support, % of stride time	IG	47.6 (0.7)	t(98)=0.733, p=0.465	46.8 (0.7)	-0.8 (0.7), [-2.2, 0.6]	t(51)=1.122, p=0.267*	F(1,302)=0.016 to 2.222, p=0.137 to 0.901	0.000 to 0.007
	CG	48.4 (1.0)		48.0 (0.9)	-0.4 (1.0), [-2.5, 1.7]	t(32)=0.382, p=0.705		
Stance phase, % of stride time	IG	73.5 (0.3)	t(79)=1.456, p=0.149*	73.2 (0.4)	-0.3 (0.4), [-1.1, 0.5]	t(37)=0.784, p=0.438*	F(1,302)=0.092 to 3.400, p=0.066 to 0.761 [a]	0.000 to 0.011
	CG	74.4 (0.5)		73.9 (0.6)	-0.5 (0.6), [-1.8, 0.8]	t(19)=0.796, p=0.436*		
Dual-task costs, naming animals (IG: n=194, KG: n=110)								
Walking speed, %	IG	-29.3 (2.5)	t(46)=-0.684, p=0.498	-26.7 (2.3)	2.6 (2.7), [-2.8, 8.0]	t(27)=-0.977, p=0.337*	F(1,302)=0.010 to 1.900, p=0.169 to 0.922 [a, b]	0.000 to 0.006
	CG	-31.7 (2.0)		-26.8 (4.0)	4.9 (3.9), [-2.9, 12.7]	t(37)=-1.249, p=0.220*		

Stride length, %	IG	-14.9 (1.7)	t(34)=-0.565, p=0.576	-10.1 (1.8)	4.8 (2.2), [0.4, 9.2]	**t(20)=-2.235, p=0.037**	F(1,302)=0.001 to 2.157, p=0.143 to 0.980 [b]	0.000 to 0.007
	CG	-16.3 (1.6)		-11.8 (2.9)	4.5 (3.0), [-1.6, 10.5]	t(22)=-1.496, p=0.149*		
Stride time, %	IG	28.2 (2.5)	t(103)=0.120, p=0.905	28.9 (2.4)	0.7 (2.8), [-4.8, 6.3]	t(63)=-0.263, p=0.794	F(1,302)=0.014 to 3.057, p=0.081 to 0.907 [a, b]	0.000 to 0.010
	CG	28.7 (2.8)		26.3 (3.9)	-2.4 (4.2), [-10.9, 6.2]	t(16)=0.561, p=0.582*		
Double support, %	IG	21.0 (1.4)	t(68)=0.089, p=0.929	17.0 (1.8)	-4.0 (2.1), [-8.1, 0.1]	t(32)=1.931, p=0.063*	F(1,302)=0.013 to 2.661, p=0.104 to 0.911	0.000 to 0.009
	CG	21.2 (1.9)		18.4 (2.6)	-2.8 (3.0), [-8.9, 3.3]	t(25)=0.923, p=0.365*		
Stance phase, %	IG	5.4 (0.4)	t(62)=1.416, p=0.162*	4.3 (0.5)	-1.1 (0.6), [-2.3, 0.1]	t(29)=1.807, p=0.081*	F(1,302)=0.000 to 3.328, p=0.069 to 0.996 [b]	0.000 to 0.011
	CG	6.4 (0.6)		5.1 (0.8)	-1.4 (0.9), [-3.1, 0.4]	t(22)=1.601, p=0.123*		

CG: control group, CI_{95}: 95 % confidence interval, df: degrees of freedom, IG: intervention group, M: mean, n: number, SE: standard error

* statistically significant in single imputations, [a] variance homogeneity not fulfilled in all imputations, [b] covariance homogeneity not fulfilled in all imputations

Statistically significant results appear bold for α=0.05. When considering adjusted significance levels using Bonferroni-Holm correction for multiple comparisons, no statistically significant results were observed.

Supplementary Table 3

***Table M.** Differences in baseline motor and cognitive performance as well as etiology of dementia and the use of walking aids between positive, non-, and negative responders in the intervention group (statistical nonsignificant results, per protocol analysis)

	Negative responders	**Non-responders**	**Positive responders**	**Between group difference**
		Mean (SD)		$F(df_{numerator}, df_{denominator})$/ $Chi^2(df)$, p
Single task, walking speed				
FICSIT-4 (n=88)	2.4 (1.3)	2.4 (1.4)	2.1 (1.7)	$Chi^2(2)=0.726$, p=0.695
Walking speed, m/sec (n=89)	0.67 (0.19)	0.69 (0.18)	0.59 (0.22)	F(2,86)=1.771, p=0.176
TUG, sec (n=89)	23.3 (12.5)	20.9 (12.5)	23.4 (10.1)	$Chi^2(2)=1.802$, p=0.406
Modified 30s CST (n=77)	7.7 (3.4)	8.7 (3.5)	7.7 (4.6)	F(2,74)=0.754, p=0.474
Modified SPPB (n=84)	6.6 (2.7)	7.2 (2.4)	6.6 (3.4)	$Chi^2(2)=0.694$, p=0.707
Clock Drawing Test (n=81)	3.1 (1.5)	3.1 (1.2)	2.7 (1.2)	F(2,78)=0.397, p=0.674
Digit Span forward (n=87)	4.5 (1.9)	5.4 (1.6)	5.1 (1.6)	F(2,84)=2.725, p=0.071
Digit Span backward (n=86)	2.4 (1.7)	2.8 (1.8)	2.6 (1.2)	$Chi^2(2)=0.555$, p=0.758
Trail Making Test (n=78)	23.3 (14.3)	20.1 (13.7)	21.5 (17.4)	$Chi^2(2)=0.783$, p=0.676
Etiology, AD % (n=63)	51.9 %	64.0 %	90.9 %	$Chi^2=5.159$, p=0.063
Walking aid, % (n=89)	71.0 %	60.5 %	73.3 %	$Chi^2(2)=1.289$, p=0.525
Single task, stride length				
FICSIT-4 (n=88)	2.3 (1.3)	2.4 (1.4)	2.0 (1.8)	$Chi^2(2)=1.315$, p=0.518
Modified 30s CST (n=77)	8.1 (3.4)	8.6 (3.9)	7.0 (2.7)	F(2,74)=0.906, p=0.409
Modified SPPB (n=84)	6.5 (2.9)	7.2 (2.5)	6.5 (2.9)	$Chi^2(2)=1.137$, p=0.567
Clock Drawing Test (n=81)	2.8 (1.5)	3.2 (1.3)	2.5 (0.9)	F(2,78)=1.671, p=0.195
Digit Span forward (n=87)	4.6 (1.8)	5.3 (1.7)	5.1 (1.9)	F(2,84)=1.238, p=0.295
Digit Span backward (n=86)	2.3 (1.6)	2.7 (1.8)	2.9 (1.5)	$Chi^2(2)=0.653$, p=0.721
Trail Making Test (n=78)	19.4 (13.7)	21.6 (14.8)	25.0 (14.5)	$Chi^2(2)=1.482$, p=0.477
Etiology, AD % (n=63)	65.0 %	56.3 %	81.8 %	$Chi^2=2.222$, p=0.324
Walking aid, % (n=89)	60.9 %	62.7 %	86.7 %	$Chi^2(2)=3.376$, p=0.185
Single task, double support				
FICSIT-4 (n=88)	2.3 (1.3)	2.4 (1.4)	1.7 (1.9)	$Chi^2(2)=1.930$, p=0.381
TUG, sec (n=89)	23.7 (10.5)	19.9 (9.7)	31.2 (22.7)	$Chi^2(2)=4.807$, p=0.090
Modified SPPB (n=84)	6.5 (2.8)	7.3 (2.4)	5.4 (3.6)	$Chi^2(2)=3.029$, p=0.220
Clock Drawing Test (n=81)	3.0 (1.4)	3.1 (1.3)	2.5 (1.2)	$Chi^2(2)=1.452$, p=0.493
Digit Span forward (n=87)	4.5 (1.7)	5.3 (1.7)	5.1 (2.0)	F(2,84)=1.862, p=0.162
Digit Span backward (n=86)	2.1 (1.6)	2.9 (1.8)	2.7 (1.4)	$Chi^2(2)=2.861$, p=0.243
Trail Making Test (n=78)	19.7 (15.2)	23.2 (13.8)	16.1 (15.4)	$Chi^2(2)=2.411$, p=0.300
Etiology, AD % (n=63)	57.9 %	63.2 %	83.3 %	$Chi^2=1.147$, p=0.626

Walking aid % (n=89)	73.1 %	61.1 %	77.8 %	Chi²=1.556, p=0.467
Dual task, counting backwards, walking speed				
FICSIT-4 (n=61)	2.8 (1.4)	2.8 (1.4)	2.2 (1.6)	Chi²(2)=2.368, p=0.306
TUG, sec (n=62)	19.4 (8.6)	17.7 (6.1)	23.1 (15.1)	Chi²(2)=2.429, p=0.297
Modified 30s CST (n=56)	8.9 (3.0)	8.1 (3.1)	8.1 (4.0)	Chi²(2)=1.667, p=0.434
Modified SPPB (n=60)	7.7 (2.2)	6.9 (2.1)	7.2 (2.9)	F(2,57)=0.544, p=0.584
Counting backwards (n=62)	17.9 (8.7)	15.4 (9.0)	17.1 (7.2)	F(2,59)=0.393, p=0.676
MMSE (n=62)	19.0 (3.7)	18.3 (3.5)	18.0 (4.5)	F(2,59)=0.384, p=0.683
Clock Drawing Test (n=58)	3.4 (1.4)	3.6 (1.0)	2.8 (0.9)	Chi²(2)=5.333, p=0.070
Digit Span forward (n=61)	5.5 (1.5)	5.0 (1.8)	5.2 (1.7)	F(2,58)=0.431, p=0.652
Digit Span backward (n=61)	3.4 (1.7)	2.4 (1.7)	2.7 (1.5)	Chi²(2)=3.065, p=0.216
Trail Making Test (n=56)	24.0 (13.5)	20.0 (15.2)	26.1 (14.2)	F(2,53)=0.762, p=0.472
Etiology, AD % (n=41)	62.5 %	36.4 %	71.4 %	Chi²=3.157, p=0.230
Walking aid, % (n=62)	62.5 %	64.3 %	70.8 %	Chi²=0.472, p=0.837
Dual task, counting backwards, stride length				
FICSIT-4 (n=61)	2.6 (1.2)	2.5 (1.5)	2.6 (1.6)	Chi²(2)=0.059, p=0.971
Stride length, cm (n=62)	82.7 (16.9)	81.8 (18.1)	69.5 (19.8)	F(2,59)=3073, p=0.054
TUG, sec (n=62)	18.2 (8.3)	20.1 (7.5)	22.4 (16.9)	Chi²(2)=0.682, p=0.711
Modified 30s CST (n=56)	8.7 (3.4)	8.5 (3.1)	8.1 (3.9)	F(2,53)=0.138, p=0.871
Modified SPPB (n=60)	7.8 (2.0)	7.0 (2.3)	7.6 (3.0)	F(2,57)=0.613, p=0.545
Counting backwards (n=62)	15.3 (9.3)	17.2 (7.4)	17.8 (8.9)	F(2,59)=0.344, p=0.710
MMSE (n=62)	18.7 (4.1)	18.1 (3.8)	18.8 (4.3)	Chi²(2)=0.706, p=0.703
Clock Drawing Test (n=58)	3.2 (0.6)	3.3 (1.4)	3.1 (1.1)	F(2,55)=0.161, p=0.852
Digit Span forward (n=61)	5.7 (2.0)	5.1 (1.6)	5.3 (1.4)	F(2,58)=0.532, p=0.590
Digit Span backward (n=61)	3.2 (1.3)	3.0 (1.9)	2.7 (1.3)	Chi²(2)=0.920, p=0.631
Trail Making Test (n=56)	22.0 (11.7)	23.0 (14.7)	26.1 (15.0)	F(2,53)=0.333, p=0.718
Etiology, AD % (n=41)	62.5 %	50.0 %	69.2 %	Chi²=1.269, p=0.546
Walking aid, % (n=62)	66.7 %	64.5 %	68.4 %	Chi²=0.149, p>0.999
Dual task, counting backwards, double support				
FICSIT-4 (n=61)	2.6 (1.2)	2.4 (1.6)	2.9 (1.5)	Chi²(2)=0.915, p=0.633
TUG, sec (n=62)	19.5 (8.2)	19.0 (7.1)	26.8 (22.2)	Chi²(2)=0.584, p=0.747
Modified 30s CST (n=56)	8.3 (3.6)	8.7 (3.4)	7.3 (2.8)	F(2,53)=0.516, p=0.600
Modified SPPB (n=60)	7.2 (2.5)	7.2 (2.5)	7.7 (2.4)	F(2,57)=0.114, p=0.893
Counting backwards (n=62)	15.9 (8.2)	17.9 (8.2)	16.2 (8.2)	F(2,59)=0.407, p=0.668
MMSE (n=62)	18.8 (3.9)	17.9 (3.9)	19.6 (4.5)	F(2,59)=0.885, p=0.418
Clock Drawing Test (n=58)	3.4 (1.4)	3.2 (1.0)	3.1 (1.4)	F(2,55)=0.304, p=0.739
Digit Span forward (n=61)	5.0 (1.7)	5.4 (1.6)	5.4 (1.7)	Chi²(2)=2.528, p=0.283
Digit Span backward (n=61)	3.1 (1.5)	2.9 (1.9)	2.9 (1.1)	Chi²(2)=0.260, p=0.878
Trail Making Test (n=56)	19.6 (12.7)	26.3 (14.3)	22.4 (15.6)	Chi²(2)=2.949, p=0.229
Etiology, AD % (n=41)	66.7 %	52.2 %	66.7 %	Chi²=0.887, p=0.676

Walking aid, % (n=62)	66.7 %	64.7 %	70.0 %	Chi²=0.150, p>0.999
Dual task, naming animals, walking speed				
FICSIT-4 (n=61)	2.7 (1.2)	2.4 (1.6)	2.5 (1.6)	Chi²(2)=0.338, p=0.844
Modified 30s CST (n=53)	9.7 (3.3)	9.2 (4.3)	7.2 (2.9)	F(2,50)=2.518, p=0.091
Modified SPPB (n=59)	7.6 (2.5)	7.2 (2.7)	6.8 (2.4)	F(2,56)=0.542, p=0.584
Verbal fluency animals (n=61)	7.9 (3.3)	8.7 (4.3)	8.1 (3.3)	F(2,58)=0.240, p=0.787
MMSE (n=61)	17.4 (4.2)	17.7 (4.4)	18.1 (4.5)	F(2,58)=0.121, p=0.886
Clock Drawing Test (n=56)	3.3 (1.6)	3.1 (1.3)	3.0 (0.8)	F(2,53)=0.203, p=0.817
Digit Span forward (n=60)	5.2 (1.3)	5.1 (1.9)	4.9 (2.2)	Chi²(2)=0.634, p=0.728
Digit Span backward (n=60)	2.6 (1.0)	3.5 (1.9)	2.4 (1.8)	Chi²(2)=5.727, p=0.057
Trail Making Test (n=56)	22.3 (15.1)	22.5 (13.0)	20.1 (14.5)	F(2,53)=0.164, p=0.849
Etiology, AD % (n=43)	71.4 %	71.4 %	53.3 %	Chi²=1.378, p=0.561
Walking aid, % (n=61)	61.9 %	63.2 %	76.2 %	Chi²(2)=1.178, p=0.555
Dual task, naming animals, stride length				
FICSIT-4 (n=61)	2.1 (1.4)	2.9 (1.4)	2.3 (1.6)	Chi²(2)=3.794, p=0.150
TUG, sec (n=61)	20.8 (9.7)	20.2 (14.3)	24.3 (13.3)	Chi²(2)=3.250, p=0.197
Modified 30s CST (n=53)	9.6 (4.1)	9.1 (3.6)	7.4 (3.2)	F(2,50)=1.486, p=0.236
Modified SPPB (n=59)	6.9 (2.4)	7.7 (2.3)	6.7 (2.8)	F(2,56)=0.944, p=0.395
Verbal fluency animals (n=61)	7.9 (3.3)	8.5 (4.1)	8.1 (3.2)	F(2,58)=0.151, p=0.860
MMSE (n=61)	15.6 (4.1)	18.2 (4.1)	18.7 (4.6)	Chi²(2)=4.724, p=0.094
Clock Drawing Test (n=56)	3.1 (1.2)	3.0 (1.4)	3.3 (1.0)	F(2,53)=0.338, p=0.715
Digit Span forward (n=60)	5.1 (1.3)	5.0 (2.0)	5.2 (1.9)	Chi²(2)=0.164, p=0.921
Digit Span backward (n=60)	2.6 (1.3)	2.9 (1.9)	2.8 (1.6)	F(2,57)=0.149, p=0.862
Trail Making Test (n=56)	19.7 (13.3)	25.0 (14.0)	17.4 (13.9)	Chi²(2)=2.758, p=0.252
Etiology, AD % (n=43)	75.0 %	58.3 %	72.7 %	Chi²=1.028, p=0.688
Dual task, naming animals, double support				
FICSIT-4 (n=61)	2.6 (1.5)	2.7 (1.4)	2.0 (1.6)	Chi²(2)=2.576, p=0.276
TUG, sec (n=61)	20.1 (10.4)	20.2 (13.2)	26.4 (14.8)	Chi²(2)=3.326, p=0.190
Modified 30s CST (n=53)	9.8 (3.6)	9.0 (3.5)	6.6 (3.6)	F(2,53)=2.483, p=0.094
Modified SPPB (n=59)	7.5 (2.7)	7.6 (2.2)	5.8 (2.9)	F(2,56)=2.366, p=0.103
Verbal fluency animals (n=61)	7.6 (2.7)	9.1 (4.0)	6.6 (2.9)	F(2,58)=2.592, p=0.084
MMSE (n=61)	16.8 (3.9)	18.3 (4.5)	17.3 (4.4)	Chi²(2)=1.889, p=0.389
Clock Drawing Test (n=56)	3.5 (1.8)	2.9 (1.2)	3.2 (0.8)	F(2,56)=0.809, p=0.451
Digit Span forward (n=60)	5.1 (1.3)	5.1 (2.1)	5.0 (1.5)	Chi²(2)=0.118, p=0.943
Digit Span backward (n=60)	2.5 (1.2)	3.2 (1.8)	2.3 (1.6)	F(2,57)=1.563, p=0.218
Trail Making Test (n=56)	21.2 (14.2)	24.3 (13.5)	15.5 (14.0)	Chi²(2)=3.505, p=0.173
Etiology, AD % (n=43)	75.0 %	62.5 %	63.6 %	Chi²=0.453, p=0.910
Walking aid, % (n=61)	71.4 %	61.8 %	76.9 %	Chi²=1.042, p=0.663

30s CST: 30-second chair stand test, AD: Alzheimer's disease, df: degrees of freedom, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, M: mean, MMSE: Mini-Mental State Examination, n: number, SD: standard deviation, SPPB: Short Physical Performance Battery, TUG: Timed Up & Go Test

Supplementary Table 4

***Table N.** Heterogeneity of sample in motor and cognitive baseline performance (per protocol)

	Total sample	**Intervention group**	**Control group**
Balance			
FICSIT-4 [Median (IQR), range]	2 (2), 0-5 (n=160)	2 (2.5), 0-5 (n=89)	2 (2), 0-4 (n=71)
Mobility			
TUG, sec [Mean (SD), range]	22.5 (11.2), 9.0-85.0 (n=160)	22.3 (12.2), 9.7-85.0 (n=90)	22.7 (10.0), 9.0-61.0 (n=70)
Lower limb strength & function			
modified 30s CST [Mean (SD), range]	8.0 (3.6), 1-19 (n=141)	8.1 (3.7), 1-19 (n=78)	7.9 (3.6), 2-17 (n=63)
SPPB [Mean (SD), range]	6.5 (2.7), 1-12 (n=152)	6.8 (2.8), 1-12 (n=85)	6.2 (2.6), 1-12 (n=67)
Executive function			
Clock Drawing Test [Mean (SD), range]	3.0 (1.5), 1-9 (n=152)	3.0 (1.3), 1-8 (n=82)	3.1 (1.7), 1-9 (n=70)
Trail Making Test [Mean (SD), range]	21.3 (13.0), 0-48 (n=144)	21.5 (14.3), 0-48 (n=79)	21.0 (11.4), 1-43 (n=65)
Attention & working memory			
Digit Span forward [Mean (SD), range]	5.0 (1.7), 1-10 (n=161)	5.0 (1.7), 1-10 (n=88)	5.0 (1.8), 1-9 (n=73)
Digit Span backward [Mean (SD), range]	2.8 (1.7), 0-6 (n=160)	2.7 (1.7), 0-6 (n=87)	2.9 (1.7), 0-6 (n=73)

30s CST: 30-second chair stand test, FICSIT-4: Frailty and Injuries: Cooperative Studies of Intervention Techniques - subtest 4, IQR: interquartile range, n: number, SD: standard deviation, SPPB: Short Physical Performance Battery, TUG: Timed Up & Go Test

Contributions to research manuscripts

The five research manuscripts in this dissertation were written within the framework of the research project Physical Activity Against Dementia.

The author of this dissertation, Sandra Trautwein, was member of the research team of this research project. She contributed to conceiving the idea and developing the conception as well as the study design for the study. Moreover, she coordinated the study and was responsible for its implementation.

Manuscripts 1, 2, 3, and 5 were written by Sandra Trautwein and arose under her full responsibility. She was responsible for the ideas of this research manuscripts, their preparation and submission, as well as responding to reviewer's comments. She contributed in data acquisition, statistical analysis, and data interpretation.

Manuscript 4 was written by Dr. Bettina Barisch-Fritz. Sandra Trautwein helped drafting the manuscript and contributed in data acquisition, statistical analysis, and data interpretation.

Acknowledgements

Viele Personen haben mich auf dem Weg meiner Promotion begleitet und unterstützt. Allen möchte ich von Herzen danken. Sie waren während meiner Promotionsphase eine wichtige Stütze für mich und haben auf vielfältigste Weise zum Gelingen meiner Promotion beigetragen.

Besonders bedanken möchte ich mich bei meinem Betreuer und Doktorvater Prof. Dr. Alexander Woll, der mich seit Beginn meines Studiums hervorragend wissenschaftlich unterstützt, gefördert und begleitet hat und mir dadurch meine Promotion ermöglichte. Durch sein Vertrauen und seinen Glauben an mich konnte ich viele wichtige Erfahrungen sammeln. Seine menschliche, verständnisvolle sowie herzliche Art und Führung haben es mir leicht gemacht, jederzeit mit Fragen und Problemen zu ihm zu kommen. Dafür schätze ich ihn sehr.

Mein herzlicher Dank gilt Prof. Dr. Thorsten Stein für die Erstellung des Zweitgutachtens. Er hat durch seine kompetente und anspruchsvolle Mitwirkung bei der Auswahl der motorischen Testverfahren sowie der Planung und Diskussion der Analyse der Gangdaten einen für mich äußerst wichtigen inhaltlichen Beitrag geleistet. In den gewinnbringenden Gesprächen während der unterschiedlichen Phasen meiner Promotion hat mich seine Wertschätzung und motivierende Unterstützung immer wieder sehr beeindruckt.

Ein riesiges Dankeschön möchte ich meinem „Demenzteam“ aussprechen. Ohne Andrea Scharpf, Luisa Appelles, Dr. Bettina Barisch-Fritz und Jelena Bezold wären unsere Projekte Bewegung gegen Demenz und InCoPE nie so möglich gewesen. Es ist etwas Besonderes, dass wir uns im Team so gut ergänzen, stets aufeinander verlassen können und immer an einem gemeinsamen Strang ziehen. Ich werde mich immer an die intensive, schöne, erlebnisreiche, stressige und prägende Zeit erinnern und nie vergessen, wie viel wir auch miteinander gelacht haben. Andy und ich haben uns zu Beginn durch eine enorm lehrreiche aber auch harte Projektphase gekämpft, in der wir gemeinsam unglaublich viele Erfahrungen gesammelt haben. Ohne sie und den gegenseitigen Zuspruch hätte ich das nicht geschafft. Unsere zahlreichen Gespräche und Telefonate waren sehr wertvoll für mich. Luisa war für mich ein kleiner Wirbelwind, der genau zum richtigen Zeitpunkt zu uns ins Projekt gekommen ist und mit vollem Tatendrang und Elan dort angepackt hat, wo immer sie gebraucht wurde. Bettina hat die für mich unglaublich wichtige Rolle als Mentorin übernommen und hatte jederzeit

(!!!) ein offenes Ohr für meine Fragen und Probleme, egal ob sie wissenschaftlicher, organisatorischer oder persönlicher Art waren. Durch ihre verständnisvollen, durchdachten und pragmatischen Ratschläge hat sie es immer wieder geschafft auf mich einzugehen und mir einen Lösungsweg zu zeigen. Jelena war durch ihr aufgeschlossenes und fröhliches Wesen in kürzester Zeit ein wichtiges Mitglied in unserem Team. Ich schätze ihre offene Art und ihre zuverlässige und strukturierte Arbeitsweise sehr. Ihr engagierter Einsatz hat es mir ermöglicht, dass ich mich vor allen in der Endphase auf meine Promotion konzentrieren konnte.

Für die gewinnbringenden Diskussionen, den konstruktiven Austausch, die Unterstützung bei der Interpretation der Ergebnisse und der Erstellung von wissenschaftlichen Publikationen danke ich Dr. Janina Krell-Rösch, Philipp Maurus und Dr. Steffen Ringhof. Ihr Input in den verschiedensten Bereichen war für mich eine große inhaltliche Bereicherung. Die zahlreichen Anregungen und Tipps haben mir immer wieder Denkanstöße gegeben und mich dadurch wissenschaftlich vorangebracht. Janina verdient sich außerdem ein großes Dankeschön für das fleißige Korrekturlesen und ihren Beitrag zum letzten Feinschliff der Arbeit.

Bedanken möchte ich mich außerdem bei allen Kollegen des IfSS. Das angenehme Arbeitsklima, die vielen intensiven Gespräche, die entstandenen Freundschaften, die emotionale Unterstützung, der lebhafte Arbeitsalltag, das gemeinsame Lachen, der positive Zuspruch und die vielen schönen Momenten waren eine Bereicherung und haben meinen Arbeitsplatz in den vergangenen Jahren zu einem wichtigen Lebensbereich werden lassen, an dem ich mich immer wohl gefühlt habe. Vor allem die Kellerkinder haben dazu beigetragen, dass der Alltag am IfSS mehr als nur ein normaler Job ist. Auch habe ich durch meine Kollegen viel Unterstützung bei meiner wissenschaftlichen Arbeit erfahren, z. B. durch statistische und methodische Beratung, Rat und Tat im Umgang mit den Finanzen und der Verwaltung, im Projektmanagement, bei der Diskussion und Auswertung der Gangdaten, im organisatorischen Bereich, auf praktischer und technischer Ebene und bei vielem mehr. DANKE an Iyas, Stefan, Bastian, Herrn Bös, Marion, Peggy, Didi, Alex, Dani, Darko, Sabine, Micha, Frau Kölmel, Jule, Wolfgang, Christina, Claudi, Claudio, Doris, Matthias, Steffi, Rita, Kathrin, die AG Woll, die AG Stein, das Doktorandenkolloquium und den Hausdienst.

Bei der Erstellung von wissenschaftlichen Publikationen, habe ich zahlreiches Feedback und fachliche Ratschläge erhalten. Dies war für mich ungemein lehrreich und hat

dazu beigetragen, dass die Publikationen immer nochmal ein bisschen besser, exakter, strukturierter, verständlicher und durchdachter geworden sind. Daher ein großes Dankeschön an alle Coautoren.

Ein so großes Projekt wie Bewegung gegen Demenz wäre ohne die Mitarbeit von zahlreichen Helfern nicht möglich gewesen. Ich bedanke mich bei allen unseren Hiwis, Praktikanten, Trainern und Testleitern für die tatkräftige Unterstützung und den unermüdlichen Einsatz. Ihr Engagement und Tatendrang haben es ermöglicht eine so große Stichprobe zu erfassen. Besonders hervorzuheben ist das fleißige Arbeiten im Hintergrund sowie das zuverlässige Mitdenken von Anela Hadzic und Corinna Wehrstein, die fast über die gesamte Dauer als Hiwis im Projekt dabei waren.

Für die Bereitstellung der finanziellen Ressourcen danke ich der Dietmar Hopp Stiftung. Ihre Förderung hat es ermöglicht das Projekt Bewegung gegen Demenz umzusetzen. Ich konnte darin für mich außerordentlich wichtige Erfahrungen im Umgang mit Demenzerkrankungen sammeln und mich auf wissenschaftlicher Ebene weiter qualifizieren. Vielen Dank an Dietmar Hopp, Henrik Westerberg und Dietmar Pfähler für die kooperative und nachhaltige Zusammenarbeit. Ebenfalls bedanken möchte ich mich bei Helmut Hoffmann, der das Projekt Bewegung gegen Demenz von Seiten der Dietmar Hopp Stiftung als Senioradvisor unterstützt hat. Er hat uns bei verschiedensten Problemstellungen auf neue Ideen gebracht und durch seine Begleitung so manchen ersten Schritt erleichtert.

Mein Dank gilt auch allen Teilnehmerinnen und Teilnehmern des Projektes Bewegung gegen Demenz sowie den kooperierenden Pflegeeinrichtungen der Altenhilfe. Es ist nicht selbstverständlich sich mit solchem Engagement an einer wissenschaftlichen Studie zu beteiligen. Vielen Dank für das Vertrauen und die Unterstützung bei der zeitintensiven organisatorischen Umsetzung unseres Projektes.

Auf emotionaler Ebene haben meine Freunde einen entscheidenden Beitrag zum Gelingen meiner Promotion geleistet. Sie haben mich auf vielfältigste Art unterstützt und mir vor allem durch ihr Zuhören, ihren Zuspruch, die ermutigenden Worte und gemeinsame Unternehmungen viel Kraft und Bestätigung auf meinem Weg gegeben. Ganz besonders bedanken möchte ich mich bei Keddy Rast, Utz Obenaus und Claire Vogel. Die vielen Telefonate und Gespräche haben mir in schwierigen Phasen einfach gutge-

tan. Es ist ein großes Geschenk solche Freunde zu haben, die mich immer unterstützen, so viel Verständnis haben und in einer so besonderen Lebensphase wie der Promotion mit einem mitfühlen können.

Zum Schluss möchte ich mich von ganzem Herzen bei den wichtigsten Menschen in meinem Leben bedanken – meiner Familie. Egal in welcher Situation, ich habe immer bedingungslose und permanente Unterstützung erhalten. Meine Großeltern Mäme und Hans waren für mich die wichtigste Motivation mich mit dem Thema Bewegung und Demenz auseinander zu setzen. Sie haben mir gezeigt, was im Leben wichtig ist und mich mit ihrer liebevollen Art durch meine Kindheit, Jugend und im Erwachsenenalter begleitet. Oma und Opa danke ich für das liebevolle Zuhause, das sie mir während meiner Zeit in Karlsruhe gegeben haben. Ich habe die gemeinsame Zeit, die vielen Gespräche und das verwöhnt werden sehr genossen. Wir haben dadurch einen gemeinsamen Alltag geschenkt bekommen in dem ich immer so angenommen wurde wie ich bin, egal ob ich einen guten oder schlechten Tag hatte. Unendlich dankbar bin ich meinen Eltern für alles was sie für mich tun und dass sie immer für mich da sind. Ihr uneingeschränkter Glaube an mich und ihre liebevolle Fürsorge begleiten mich seit ich denken kann. Vor allem in schwierigen Phasen kann ich mich fallen lassen und weiß, dass ich mich mit jedem Problem an sie wenden kann. Mama und Papa haben mich immer ermutigt meinen eigenen Weg zu gehen und mich beispielslos dabei unterstützt meine Ziele zu erreichen. Mein größter Dank gilt meinem Mann Christoph, der die Zeit meiner Promotion am nächsten miterlebt und die vielen Herausforderungen gemeinsam mit mir gemeistert hat. Seine unendliche Liebe, sein tiefes Vertrauen, sein Verständnis, seine grenzenlose Fürsorge und der nicht in Worte zu fassende Rückhalt haben mich durch diese nicht immer einfache Zeit getragen. Er hat mich einfach in allem unglaublich unterstützt und immer an mich geglaubt. Tausend Dank dafür!