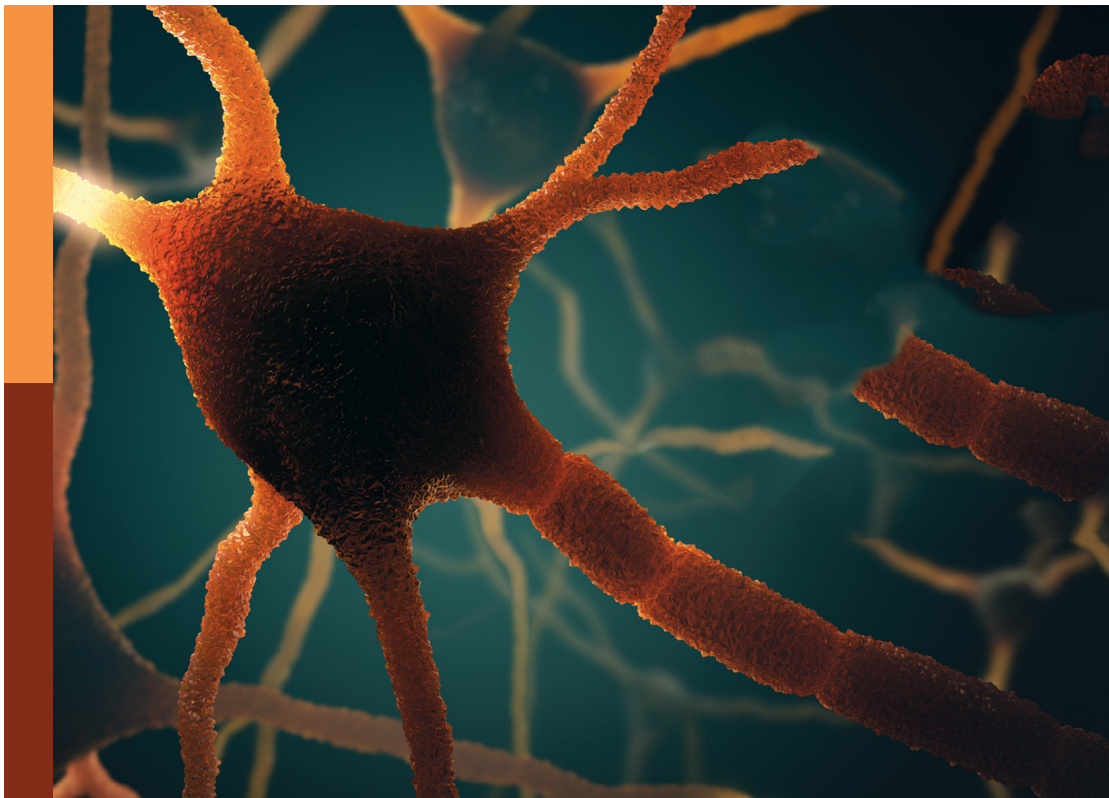
frontiers | Research Topics

Activities of daily living and everyday functioning: From normal aging to neurodegenerative diseases

Edited by
Ondrej Bezdicek, Inga Liepelt-Scarfone, Joaquim Ferreira and Robert Fellows

Published in
Frontiers in Aging Neuroscience



ISSN 1664-8714
ISBN 978-2-83252-077-2
DOI 10.3389/978-2-83252-077-2

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Activities of daily living and everyday functioning: From normal aging to neurodegenerative diseases

Topic editors

Ondrej Bezdicek — Charles University, Czechia
Inga Liepelt-Scarfone — University of Tübingen, Germany
Joaquim Ferreira — Universidade de Lisboa, Portugal
Robert Fellows — University of California, San Diego, United States

Citation

Bezdicek, O., Liepelt-Scarfone, I., Ferreira, J., Fellows, R., eds. (2023). *Activities of daily living and everyday functioning: From normal aging to neurodegenerative diseases.* Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83252-077-2

Table of contents

frontiers | Frontiers in Aging Neuroscience

TYPE Editorial
PUBLISHED 20 March 2023
DOI 10.3389/fnagi.2023.1161736

OPEN ACCESS

EDITED AND REVIEWED BY
Kristy A. Nielson,
Marquette University, United States

*CORRESPONDENCE
Ondrej Bezdicek
✉ Ondrej.Bezdicek@lf1.cuni.cz

SPECIALTY SECTION
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 08 February 2023
ACCEPTED 27 February 2023
PUBLISHED 20 March 2023

CITATION
Bezdicek O, Ferreira J, Fellows R and Liepelt-Scarfone I (2023) Editorial: Activities of daily living and everyday functioning: From normal aging to neurodegenerative diseases. *Front. Aging Neurosci.* 15:1161736. doi: 10.3389/fnagi.2023.1161736

Editorial: Activities of daily living and everyday functioning: From normal aging to neurodegenerative diseases

Ondrej Bezdicek [1]*, Joaquim Ferreira[2], Robert Fellows[3] and Inga Liepelt-Scarfone[4,5]

[1]Department of Neurology and Centre of Clinical Neuroscience, First Faculty of Medicine and General University Hospital in Prague, Charles University, Prague, Czechia, [2]Laboratory of Clinical Pharmacology and Therapeutics, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal, [3]Department of Psychiatry, University of California, San Diego, San Diego, CA, United States, [4]German Center for Neurodegenerative Diseases (DZNE) and Hertie Institute for Clinical Brain Research, Eberhard Karls University Tübingen, Tübingen, Germany, [5]IB-Hochschule, Stuttgart, Germany

KEYWORDS

activities of daily living, cognition, mild cognitive impairment, neurodegeneration, functional ability

Editorial on the Research Topic
Activities of daily living and everyday functioning: From normal aging to neurodegenerative diseases

Activities of daily living (ADLs) are a wide-ranging psychological construct to designate and measure individuals' independence and range of everyday functioning in a social environment. Populations in developed countries are aging and estimates of those over the age of 85 who will need assistance with ADLs have tripled. An even steeper decline in ADLs can be expected in people at risk of developing neurodegenerative or other types of diseases (Buchman et al., 2009; Gill et al., 2013). Deficits in ADLs due to cognitive impairment are increasingly discussed as a risk factor for dementia in older age and neurodegenerative diseases (Verlinden et al., 2016; Becker et al., 2022; Jang et al., 2022; Ng et al., 2022), indicating a need for sensitive assessments, especially in mild stages of ADL difficulty. ADLs are affected by multiple factors such as physical fitness, cognitive ability, and mental health (Cahn et al., 1998; Tam et al., 2008; Christensen et al., 2013). Thus, basic and applied research on ADLs is highly necessary as any psychological, medical, or technological intervention that helps to promote or preserve ADLs in later adulthood is of the utmost importance to preserve social independence and goes far beyond this: it helps to diminish the caregiver burden in each family, it decreases healthcare costs, and it creates an existential perspective that life is worth living.

Thus, the current *Research Topic on Activities of Daily Living and Everyday Functioning: From Normal Aging to Neurodegenerative Diseases* is dedicated to different areas of basic and applied research.

Gait disturbances and falls constitute a major reason for disability in older age as fall incidence can result in fractures and other injuries (Lord and Close, 2018). Cerebral small vessel disease (CSVD) has been related to gait disturbances (Markus and De Leeuw, 2023); however, the role of enlarged perivascular spaces in the basal ganglia on these disabilities is still in debate. Data from Yang S. et al. identified that a greater number of enlarged

perivascular spaces in the basal ganglia were independently related to gait disturbances in older people with CSVD. Fear of falling (FoF) is associated with poorer physical and cognitive functions and decreasing ability to perform ADLs; it has been identified as a potential risk factor for falls in older age and neurodegenerative diseases (Bryant et al., 2013; Schoene et al., 2019). The article by Atrsaei et al. evaluates the influence of FoF on mobility in patients with Parkinson's disease (PD), highlighting the importance of monitoring different environments and assessment strategies. In a meta-analysis, Kim et al. found that task-specific reactive balance exercise training may be an optimal intervention in preserving reactive balance to prevent falls in older age. Taken together, these studies identify brain morphological and multimodal clinical risk factors for decreased mobility, as well as targeted strategies for reducing falls and improving functioning. All contributions underline the importance of vascular, psychological, and prophylactic factors in falls and in their monitoring for a better prediction of treatment strategies.

PD is a multicomplex neurodegenerative disease comprising both motor and non-motor symptoms affecting ADLs. For the identification of ADLs impairment indicative of dementia in PD, valid ratings of motor and cognitive sources primarily affecting ADLs are essential (Dubois et al., 2007). The article by Becker et al. focuses on the agreement between self- and informant-reported ADLs and their association with cognitive performance in patients with PD. Of note, motor severity showed a high impact on both self- and informant-reported functioning. Their research indicates a need for objective ADLs measures because the agreement of patient and informant ratings of ADLs function showed only moderate agreement. The contribution of Rehman et al. applies modern machine learning methods to the analysis of gait in PD and controls in real-world and laboratory characteristics. Another line of research in PD is dedicated to non-motor effects and ADLs in patients undergoing deep brain stimulation of the subthalamic nucleus. The study by Bezdicek et al. shows a positive effect of the treatment on instrumental ADLs in the post-surgery phase. In sum, these contributions highlight novel approaches to evaluate the validity of assessments and monitor the effects of treatment on ADLs in neurodegenerative diseases such as PD.

Psychiatric diseases are known to affect patients' behavior and everyday life. Depression is common in older age, especially in patients with mild cognitive impairment (MCI; Ismail et al., 2017). The article by Numbers et al. revealed that patients' self- and caregiver ratings of ADLs are associated with the severity of depression in a community-dwelling older group. In contrast, objective measures seem to be more robust against the influence of depression and personality and might therefore be a suitable alternative to differentiate between cognitive and psychiatric effects on ADLs disabilities. Similarly, Yang D. et al. show that a certain time and proportion per week of vigorous to moderate physical activity in men over 45 years of age lowers the risk of depression. These studies emphasize the importance of preventive measures in preserving ADLs in relation to neuropsychiatric symptoms such as depression.

In addition to the identification of risk and modulating factors of ADLs in older age and neurodegenerative diseases, treatment strategies are essential to prevent the progression of ADLs impairment. In their systematic review and meta-analysis, Han et al. concluded that combined cognitive and physical intervention enhances cognition in older adults in the short term irrespective of patients' cognitive status. However, to get insight into long-term treatment effects in older adults, additional high-quality studies are needed. The effects of multimodal exercise on health outcomes in community-dwelling older adults are the focus of the article of Vogel et al.. Their data support the assumption that the training outcome depends on factors like sleep duration, movement biography, and activity profile. Therefore, the identification of factors maximizing the treatment outcome of specific therapies is crucial for patients' differential treatment indication.

To date, the pathological mechanisms leading to ADLs impairment are only partly understood. The work of Fellows et al. concluded that larger white matter hyperintensity and smaller hippocampal volumes are correlates for poorer everyday function. They also identified unique contributions of cognitive measures to a newly developed index of pathological functional impairment and neuropsychiatric symptoms to functional reserve in MCI and healthy older adults. Frailty is associated with lower health-related quality of life (Solfrizzi et al., 2019; Chen et al., 2022) and increases the risk of disability, dementia, and mortality (Yi and Yoon, 2023). Most interestingly, abnormalities in the integrity of the left anterior thalamic radiation seem to be associated with frailty in patients with cardiometabolic diseases as shown by Tamura et al.. Also, the prevalence of MCI, which is a risk factor for lower ADLs functioning (Perneczk et al., 2006), according to the study of Liu et al., is higher in women than in men of older age. This group of studies uncovers the pathological and demographic factors at play in the development of ADLs impairment.

In conclusion, the Research Topic gives an overview of the state-of-the-art research on ADLs in healthy aging and different clinical conditions. It highlights the key areas of contemporary research into ADLs so that individuals can participate in social life events even at an older age.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

This work was funded by the National Institute for Neurological Research, Czech Republic, Programme EXCELES (ID Project No. LX22NPO5107) and the European Union - Next Generation EU.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Becker, S., Bode, M., Brockmann, K., Gasser, T., Michaelis, K., Solbrig, S., et al. (2022). Cognitive-driven activities of daily living impairment as a predictor for dementia in Parkinson disease: a longitudinal cohort study. *Neurology* 99, e2548–e2560. doi: 10.1212/WNL.0000000000201201

Bryant, M. S., Rintala, D. H., Hou, J.-G., and Protas, E. J. (2013). Relationship of falls and fear of falling to activity limitations and physical inactivity in Parkinson's disease. *J. Aging Phys. Activ.* 23, 187–193. doi: 10.1123/japa.2013-0244

Buchman, A. S., Boyle, P. A., Wilson, R. S., Fleischman, D. A., Leurgans, S., and Bennett, D. A. (2009). Association between late-life social activity and motor decline in older adults. *Arch. Intern. Med.* 169, 1139–1146. doi: 10.1001/archinternmed.2009.135

Cahn, D. A., Sullivan, E. V., Shear, P. K., Pfefferbaum, A., Heit, G., and Silverberg, G. (1998). Differential contributions of cognitive and motor component processes to physical and instrumental activities of daily living in Parkinson's disease. *Arch. Clin. Neuropsychol.* 13, 575–583. doi: 10.1016/S0887-6177(98)00024-9

Chen, B., Wang, M., He, Q., Wang, Y., Lai, X., Chen, H., et al. (2022). Impact of frailty, mild cognitive impairment and cognitive frailty on adverse health outcomes among community-dwelling older adults: a systematic review and meta-analysis. *Frontiers in Medicine (Lausanne)* 9, 1009794. doi: 10.3389/fmed.2022.1009794

Christensen, K., Thinggaard, M., Oksuzyan, A., Steenstrup, T., Andersen-Ranberg, K., Jeune, B., et al. (2013). Physical and cognitive functioning of people older than 90 years: a comparison of two Danish cohorts born 10 years apart. *Lancet* 382, 1507–1513. doi: 10.1016/S0140-6736(13)60777-1

Dubois, B., Burn, D., Goetz, C., Aarsland, D., Brown, R. G., Broe, G. A., et al. (2007). Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder societytask force. *Mov. Disord.* 22, 2314–2324. doi: 10.1002/mds.21844

Gill, D. P., Hubbard, R. A., Koepsell, T. D., Borrie, M. J., Petrella, R. J., Knopman, D. S., et al. (2013). Differences in rate of functional decline across three dementia types. *Alzheimer's Dementia* 9, S63–S71. doi: 10.1016/j.jalz.2012.10.010

Ismail, Z., Elbayoumi, H., Fischer, C. E., Hogan, D. B., Millikin, C. P., Schweizer, T., et al. (2017). Prevalence of depression in patients with mild cognitive impairment: a systematic review and meta-analysis. *JAMA Psychiatry* 74, 58–67. doi: 10.1001/jamapsychiatry.2016.3162

Jang, S., Numbers, K., Lam, B. C. P., Sachdev, P. S., Brodaty, H., and Reppermund, S. (2022). Performance-based vs informant-reported instrumental activities of daily living in predicting dementia. *JAMDA* 23, 1342–1347. doi: 10.1016/j.jamda.2021.09.020

Lord, S. R., and Close, J. C. T. (2018). New horizons in falls prevention. *Age Ageing* 47, 492–498. doi: 10.1093/ageing/afy059

Markus, H. S., and De Leeuw, F. E. (2023). Cerebral small vessel disease: recent advances and future directions. *Int. J. Stroke* 18, 4–14. doi: 10.1177/17474930221144911

Ng, T. P., Lee, T. S., Lim, W. S., Chong, M. S., Yap, P., Cheong, C. Y., et al. (2022). Functional mobility decline and incident mild cognitive impairment and early dementia in community-dwelling older adults: the Singapore Longitudinal Ageing Study. *Age Ageing* 51, afac182. doi: 10.1093/ageing/afac182

Perneczk, y. R., Pohl, C., Sorg, C., Hartmann, J., Tosic, N., Grimmer, T., et al. (2006). Impairment of activities of daily living requiring memory or complex reasoning as part of the MCI syndrome. *Int. J. Geriatr. Psychiatry*. 21, 158–162. doi: 10.1002/gps.1444 PMID: 16416470.

Schoene, D., Heller, C., Aung, Y. N., Sieber, C. C., Kemmler, W., and Freiberger, E. (2019). A systematic review on the influence of fear of falling on quality of life in older people: is there a role for falls? *Clin. Interv. Aging* 14, 701–719. doi: 10.2147/CIA.S197857

Solfrizzi, V., Scafato, E., Lozupone, M., Seripa, D., Schilardi, A., Custodero, C., et al. (2019). Biopsychosocial frailty and the risk of incident dementia: the Italian longitudinal study on aging. *Alzheimer's Dement.* 15, 1019–1028. doi: 10.1016/j.jalz.2019.04.013

Tam, C. W. C., Lam, L. C. W., Lui, V. W. C., Chan, W. C., Chan, S. S. M., Chiu, H. F. K., et al. (2008). Clinical correlates of functional performance in community-dwelling Chinese older persons with mild cognitive impairment. *Int. Psychogeriatr.* 20, 1059–1070. doi: 10.1017/S1041610208007345

Verlinden, V. J. A., Van Der Geest, J. N., De Bruijn, R. F. G., Hofman, A., Koudstaal, P. J., et al. (2016). Trajectories of decline in cognition and daily functioning in preclinical dementia. *Alzheimer's Dement.* 12, 144–153. doi: 10.1016/j.jalz.2015.08.001

Yi, J., and Yoon, J. Y. (2023). Cognitive frailty increases the risk of long-term care dependency in community-dwelling older adults: a nationwide cohort study in South Korea. *Geriatr. Gerontol. Int.* 23, 117–123. doi: 10.1111/ggi.14534

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 22 October 2021
doi: 10.3389/fnagi.2021.722799

Multimodal Exercise Effects in Older Adults Depend on Sleep, Movement Biography, and Habitual Physical Activity: A Randomized Controlled Trial

Oliver Vogel, Daniel Niederer and Lutz Vogt*

Department of Sports Medicine and Exercise Physiology, Institute of Sports Sciences, Johann Wolfgang Goethe-University, Frankfurt, Germany

OPEN ACCESS

Edited by:
Inga Liepelt-Scarfone, Hertie Institute for Clinical Brain Research (HIH), Germany

Reviewed by:
David W. Russ, University of South Florida, United States
Ann-Kristin Folkerts, University Hospital of Cologne, Germany
Hannah Dorothea Loenneker, University of Tuebingen, Germany

***Correspondence:**
Oliver Vogel
vogel@sport.uni-frankfurt.de

Received: *09 June 2021*
Accepted: *21 September 2021*
Published: *22 October 2021*

Citation:
Vogel O, Niederer D and Vogt L (2021) Multimodal Exercise Effects in Older Adults Depend on Sleep, Movement Biography, and Habitual Physical Activity: A Randomized Controlled Trial. Front. Aging Neurosci. 13:722799. doi: 10.3389/fnagi.2021.722799

Background: The promotion of healthy aging is one of the major challenges for healthcare systems in current times. The present study investigates the effects of a standardized physical activity intervention for older adults on cognitive capacity, self-reported health, fear of falls, balance, leg strength and gait under consideration of movement biography, sleep duration, and current activity behavior.

Methods: This single-blinded, randomized controlled trial included 49 community-dwelling older adults (36 women; 82.9 ± 4.5 years of age (Mean [M] ± SD); intervention group = 25; control group = 24). Movement biography, sleep duration, cognitive capacity, self-reported health status, and fear of falls were assessed by means of questionnaires. Leg strength, gait, and current activity levels were captured using a pressure plate, accelerometers, and conducting the functional-reach and chair-rising-test. The multicomponent intervention took place twice a week for 45 min and lasted 16 weeks. Sub-cohorts of different sleep duration were formed to distinguish between intervention effects and benefits of healthy sleep durations. Change scores were evaluated in univariate analyses of covariances (ANCOVAs) between groups and sub-cohorts of different sleep duration in both groups. Changes in cognitive capacity, self-reported health, fear of falls, balance, leg strength, and gait were investigated using the respective baseline values, movement biography, and current activity levels as covariates. Analysis was by intention-to-treat (ITT).

Results: We found sub-cohort differences in cognitive capacity change scores [$F_{(3,48)} = 5.498$, $p = 0.003$, $\eta p^2 = 0.287$]. Effects on fear of falls [$F_{(1,48)} = 12.961$, $p = 0.001$, $\eta p^2 = 0.240$] and balance change scores $F_{(1,48)} = 4.521$, $p = 0.040$, $\eta p^2 = (0.099)$ were modified by the level of current activity. Effects on gait cadence were modified by the movement biography [$F_{(1,48)} = 4.545$; $p = 0.039$, $\eta p^2 = 0.100$].

Conclusions: Unlike for functional outcomes, our multicomponent intervention in combination with adequate sleep duration appears to provide combinable beneficial effects for cognitive capacity in older adults. Trainability of gait, fear of falls, and flexibility seems to be affected by movement biography and current physical activity levels.

Trial registration: This study was registered at the DRKS (German Clinical Trials Register) on November 11, 2020 with the corresponding trial number: DRKS00020472.

Keywords: sleep duration, cognitive performance, active aging, accelerometry, gait performance

BACKGROUND

Following the demographic change and the herewith associated increase in civilization disease incidences, healthy aging has increasingly become a key responsibility of modern healthcare systems (Friedman et al., 2019). A sufficient amount and quality of sleep as well as current and lifetime physical activity are major contributors to healthy aging (Reid et al., 2006; Daskalopoulou et al., 2017; Gopinath et al., 2018).

Physical activity in the form of complying with guidelines, maintaining lifelong activity, and exercising is known to contribute to several indicators of health (Haskell et al., 2009; Piercy and Troiano, 2018; Ruegsegger and Booth, 2018). Health parameters affected by physical activity cover a wide variety, for instance, including an increase in functional capacity as well as the risk reduction of various diseases (Vlietstra et al., 2018; Posadzki et al., 2020). Health benefits associated with sleep behavior comprise, *inter alia*, cardiometabolic health, and adiposity (Chaput et al., 2020).

However, several sleep parameters as well as markers of habitual physical activity behavior are reported to deteriorate along with the aging process (Ohayon et al., 2004; Speakman and Westerterp, 2010). These age-related declines in physical activity and sleep quality have been associated with losses in cognitive and functional capacity (Gadie et al., 2017; Landi et al., 2018; Trieu and Alessi, 2019; Chaput et al., 2020). Consequently, the maintenance of physical activity levels and sleep quality contributes to healthy aging.

Besides general physical activity and sleep behavior, interventions integrating exercise (e.g., resistance or endurance training), motor-cognitive (e.g., dual-task walking, attention, and memory tasks), and mental [e.g., relaxation techniques or mindfulness based stress reduction (MBSR)] components promote healthy aging in older adults (Fessel et al., 2017; Jadczak et al., 2018). Such multicomponent interventions improved cognitive dimensions (executive function, memory, and working memory), health-related self-report measures (quality of life, fear of falling), and functional outcomes (postural control, gait, strength) (Cadore et al., 2013; Freiberger et al., 2013; Knobf et al., 2014; Tarazona-Santabalbina et al., 2016; Fessel et al., 2017; Northey et al., 2018). Further, dimensions of sleep like duration and efficiency were enhanced (Mendelson et al., 2016; Dolezal et al., 2017).

Besides direct effects on healthy aging, most health-related factors exhibit further indirect effects and mutual relationships. To date, evidence on the interaction between physical activity and sleep in humans is scarce. However, mutual effects evoked by growth hormones and signaling molecules [brain derived neurotrophic factor (BDNF) and interleukin-6 (IL-6)] are indicated (Tan et al., 2020). Consequently, beyond being an effector or outcome in intervention studies, sleep duration is related to physical activity (McClain et al., 2014; Mendelson et al., 2016). In addition, old age habitual physical activity is predicted by movement biography (Vogel et al., 2021a). According to that, sleep parameters and physical activities are not exclusively related to healthy aging, but exhibit mutual interactions and assumingly respond to multicomponent interventions. Sleep duration, current and lifetime physical activity may consequently modify exercise intervention effects. However, to our knowledge, no study evaluated their relation to, or impact on, exercise intervention effects. An increase in the effects and efficacy of multimodal exercise interventions might contribute to the promotion of healthy aging.

The present study investigates the effects of a multicomponent exercise intervention for older adults stratified by sleep duration. We hypothesize that (1) the exercise intervention enhances cognitive and functional outcomes in comparison to controls and (2) effects differ by sleep duration.

METHODS

Study Design

This single-blinded, randomized controlled trial was approved by a local review board (Goethe-University, Department of Psychology and Sports Sciences, Ethics Committee, Approval number: 2019-22) and conducted in accordance with the Declaration of Helsinki. All participants signed informed consent before inclusion.

Participants

Inclusion criteria comprised a minimum age of 65 years, the capability of walking, and a community dwelling housing situation in a retirement home, combined with a self-reliant lifestyle. Exclusion criteria consist of dementia [Montreal cognitive assessment (MoCA)] <17 points (Freitas et al., 2018) and acute injuries or infections.

Recruitment was conducted at the respective residences of participants by personal contact subsequently to presentations of the study at meetings of residents. Measurements and

intervention units were conducted at the respective residences of participants as well. The participants were recruited in three different institutions located in Frankfurt (Hessen, Germany) and surrounding areas.

The sample size was determined using G*Power (version 3.1). The calculation was based on cognitive capacity as primary outcome, assuming an α-error of 5%, β-error of 20%, and effect size of $F = 0.235$ (Papp et al., 2009; Mewborn et al., 2017; Gheysen et al., 2018). The sample size was determined for a fixed effects ANCOVA regarding main effects and interactions. The sample size calculation yielded a required total of 43 participants. Assuming a 20% dropout-rate, 52 participants were recruited (Chatfield et al., 2005).

Experimental Setup

Block randomization was conducted to create two equal groups. Each participant was randomly assigned to either an intervention or control group (CG). Blinded allocation sequence compiling, participant enrollment, and participant assignment were performed by the director of studies. The allocation sequence for randomization was created using BiAS (BiAS for Windows, Frankfurt, Germany). Assessors were blinded to group allocation of the participants; the director of studies was not involved in the conduction of the intervention or assessments.

Contents of the Multimodal Intervention

The intervention group (IG) received a multicomponent training, based on a previously published training protocol (Cordes et al., 2019). Frequency, intensity, and duration of the training exercises were adapted on a group level with regard to the assumingly higher performance in the investigated cohort of community-dwelling older adults. The intervention was conducted by a sports scientist, following a standardized manual. Training frequency was twice a week for 16 weeks, duration was 45 min. The exercise units were structured in (1) warm-up, (2) walking tasks, (3) sitting gymnastics, and (4) flexibility and relaxation exercises. The difficulty of exercises was increased over the 16 weeks by switching task complexity, increasing repetitions, or adding weights (**Table 1**). The progression of the difficulty of exercises was based on the perceived exertion of participants (Williams, 2017). Rating of perceived exertion (RPE) was conducted every 4 weeks to check if the participants are capable of the next increment of task complexity, intensity, and/or duration. An average RPE of <12 was used as an indicator to increase the level of difficulty.

The warm-up and cool-down phase took 5 to 10 min each. The major part consisted of walking tasks and sitting gymnastics, lasting 15–20 mins each. Warm-up comprised movement games aiming to prepare the participants for the physical demands of the respective unit and relax the atmosphere by utilizing group-based games encouraging interaction. Walking tasks provided for endurance training, while simultaneous processing of dual tasks targeted cognitive capacity and balance. Walking distance was set to 15 m consistently, whereas, repetitions progressed from 15 to 25 walks per training session throughout the intervention period. Sitting gymnastics were utilized as the basis for strengthening exercises with intermittent coordination tasks serving as breaks in between physically strenuous exercises. Strengthening exercises progressed from 10 to 20 repetitions, from two to three sets and no additional weight to 1 or 2 kg. Balancing tasks progressed in the form of switching body position (e.g., sitting, standing, standing on one leg) to a higher level of difficulty throughout the intervention period. Cool-down comprised stretching for flexibility and ending the exhaustive exercises. The relaxation part was supposed to increase physical and mental well-being. The intervention was instructed as group course of maximum 10 participants.

The waiting-list CG received no additional training and was told to keep up their habitual activity levels during the intervention period.

Effect Estimator Outcomes

The effect estimator outcomes were assessed by sports scientists prior to and immediately following the 16-week intervention. The intervention and measurements were conducted by different sports scientists to assure blinding. The outcomes comprised cognitive capacity, subjective health status, fear of falling, leg strength, balance, and gait parameters.

Cognitive capacity, as the primary outcome, was captured by means of MoCA; a 12-item screening tool for executive function, visuospatial abilities, language, attention, concentration and working memory, abstract reasoning, memory, and orientation. On a scale of 0 to 30 points, a higher value represents a better cognitive capacity. We utilized the total score for further processing. The MoCA shows satisfactory psychometric properties in older adults (Thomann et al., 2020).

Self-reported health status was captured by means of the Shortform-12 (SF-12), a reduced version of the SF-36. Asking for physical and mental health in 12 items, the questionnaire shows adequate psychometric properties (Drixler et al., 2020). The SF-12 yields a mental and physical score we utilized for further processing. Higher scores (mental, physical) represent better self-rated health.

Falls self-efficacy was captured by means of the German version of the abbreviated Falls Efficacy Scale (Short FES-I). The version was reduced from 16 to 7 items for concerns of falling in different situations. A higher total score value represents more concerns of falling. We utilized the total score for further analyses. Psychometric properties of the short version reach satisfactory levels in older adults (Kempen et al., 2007, 2008). The outcome of the FES-I questionnaire is often referred to as "fear of falling" in the literature. In the following, we adopt the wording of "fear of falling," under consideration of the original wording of the outcome (fall-related self-efficacy). Despite its interchangeable use in the literature and its relation, we explicitly state that fear of falls and falls self-efficacy are not the same. All questionnaires were completed by, or with the help of, assessors in form of interviews.

For leg strength rating, assessed by the chair-rising-test, the participants were asked to stand up from a chair five times in a row with their arms crossed in front of their chest. The time required to complete the task was measured by the assessor using a standard stopwatch. A shorter time to complete the test indicates higher levels of leg strength. The chair-rising-test

TABLE 1 | Intervention exercises.

Category	Exercise	Intensity				Duration/repetitions			
		Week 1–4	**Week 5–8**	**Week 9–12**	**Week 13–16**	**Week 1–4**	**Week 5–8**	**Week 9–12**	**Week 13–16**
Balance	Romberg-stance	X	-	-	-	30–60 s	-	-	-
	Semi-tandem stance	-	X	-	-	-	30–60 s	-	-
	Tandem stance	-	-	X		-	-	30–60 s	-
	Single leg stance	-	-	-	X	-	-	-	30–60 s
	Shifting weight between heels and toes	X	X	X	X	30–60 s	30–60 s	30–60 s	30–60 s
Coordination/ cognition	Throwing and catching chiffon cloths	Unilateral	Bilateral	Contralateral	Contralateral	3 min	3 min	3 min	3 min
	Throwing and passing balls	Both hands/One ball	One-handed/One ball	Both hands/One ball	One-handed/Multiple balls	4 min	4 min	4 min	4 min
	Single leg stance	No additional task	Drawing figures with the free leg	Drawing figures with the free leg	Drawing figures with the free leg	4 min	4 min	4 min	4 min
	Memory tasks	Variable complexity	Variable complexity	Variable complexity	Variable complexity	5 min	5 min	5 min	5 min
Gait/endurance	Walking at preferred speed	Normal step size/No additional obstacles	Large step size/No additional obstacles	Normal step size/Additional obstacles	Large step size/Additional obstacles	5–10 min	5–10 min	5–10 min	5–10 min
	Fast walking	Normal step size/No additional obstacles	Large step size/No additional obstacles	Normal step size/Additional obstacles	Large step size/Additional obstacles	5–10 min	5–10 min	5–10 min	5–10 min
	Intermitted walking	Sitting and standing up task every 20 m	Sitting and standing up task every 20 m	Sitting and standing up task every 10 m	Sitting and standing up task every 10 m	3–5 min	3–5 min	3–5 min	3–5 min
	Intermitted walking	Stopping and starting at signal	Stopping and starting at signal	Stopping and starting at signal	Stopping and starting at signal	3–5 min	3–5 min	3–5 min	3–5 min
	Dual task walking	-	Recognizing or reacting to signs	Recognizing or reacting to signs	Recognizing or reacting to signs	-	3–5 min	3–5 min	3–5 min

(Continued)

TABLE 1 | Continued

Category	Exercise	Intensity				Duration/repetitions			
		Week 1–4	**Week 5–8**	**Week 9–12**	**Week 13–16**	**Week 1–4**	**Week 5–8**	**Week 9–12**	**Week 13–16**
Strengthening	Knee extension and flexion (sitting)	Bodyweight	1 kg weight bands	2 kg weight bands	2 kg weight bands	3 × 15–20 repetitions	2 × 15–20 repetitions	1 × 15-20 repetitions	2 × 15–20 repetitions
	Upright rowing	Empty bar	Empty bar	Empty bar	Empty bar	1 x 15 repetitions	1 x 15 repetitions	2 x 15 repetitions	2 x 15 repetitions
	Upper body rotation	X	X	X	X	1 x 10 repetitions to each side	1 x 10 repetitions to each side	2 x 10 repetitions to each side	2 x 10 repetitions to each side
	Compressing and uncompressing the bar held horizontal in front of the body	Subjective exhaustion	Subjective exhaustion	Subjective exhaustion	Subjective exhaustion	30 s each direction	30 s each direction	60 s each direction	60 s each direction
	Compressing a towel with knees/ arms	X	X	X	X	10 x 3 s	15 x 3 s	10 x 5 s	15 x 5 s
	Pulling on a towel held in front of the body	X	X	X	X	10 x 3 s	15 x 3 s	10 x 5 s	15 x 5 s
	Lifting Legs (Sitting)	Variable lifting height	Variable lifting height	Variable lifting height	Variable lifting height	4 min	4 min	4 min	4 min
	Clenching fists and tensing arm muscles	X	X	X	X	15 x 5 s	15 x 5 s	15 x 5 s	15 x 5 s
	Standing up (no arms)	X	X	X	X	10 repetitions	15 repetitions	20 repetitions	20 repetitions
	Biceps curls	2 kg	3 kg	4 kg	5 kg	3 x 10 repetitions	3 x 10 repetitions	3 x 10 repetitions	3 x 10 repetitions
	Reaching above head	1 kg	1 kg	2 kg	2 kg	3 x 10 repetitions	3 x 10 repetitions	3 x 10 repetitions	3 x 10 repetitions
	Front raises (straight arm)	Bodyweight	Bodyweight	1 kg	1 kg	3 x 10 repetitions	3x 10 repetitions	3 x 10 repetitions	3 x 10 repetitions
Stretching/ Cool-down	Progressive muscle-relaxation	X	X	X	X	5–10 min	5–10 min	5–10 min	5–10 min
	Porcupine ball self massage	X	X	X	X	5–10 min	5–10 min	5–10 min	5–10 min
	Imaginary journey	X	X	X	X	5–10 min	5–10 min	5–10 min	5–10 min
	Stretching of the strained muscles	X	X	X	X	5–10 min	5–10 min	5–10 min	5–10 min

X, Exercise was conducted; Exercise was not conducted; kg, Kilogramm.

exhibits excellent reliability values and is confirmed as a safe test protocol in older adults (Melo et al., 2019; Mehmet et al., 2020).

The functional-reach is a recommended motor assessment, exhibiting adequate psychometric properties (Trautwein et al., 2019; Arora et al., 2020). We conducted the sitting version of the test to minimize the risk of falls in our cohort of older adults. Hence, the starting position was sitting on a chair in a neutral body position, arms stretched to the front. From the starting position, the participants were asked to reach forward as far as possible without losing contact with the sitting surface of the chair. The distance between starting and maximum reach position was measured by means of a commercial measuring tape. A further reach indicates higher levels of balance.

Spatiotemporal gait parameters (step-length, cadence, speed, double support phase, and walk-ratio) were recorded by a capacitive force-measuring platform (WinFDM v0.0.411, Zebris GmbH, Isny, Germany). Step-length is measured as the distance from toe to contralateral toe on a centered line describing the direction of movement. Cadence was calculated as steps per minute, gait speed was reported as distance (meters) walked per time (seconds). Double support phase described the percentage duration of a gait cycle with both feet touching the ground. The walk-ratio is calculated as the quotient of step-length and cadence. Adequate validity has recently been proved for the utilized platform (Rudisch et al., 2021). Participants were asked to walk at a comfortable, self-chosen speed along a 6-meter ground-level walkway containing the measuring platform. For valid results, participants crossed the walkway three times.

Effect Modifiers

The potential effect modifiers were gathered prior to the commencement of the intervention period. Movement biography, sleep duration, and current activity levels were assessed.

Movement biography was captured by means of the Lifetime Leisure Physical Activity Questionnaire (LLPAQ). The LLPAQ captures lifespan physical activity from birth up to 95 years of age. The questionnaire records time spent in different activities over lifetime divided into seven periods. The questionnaire distinguishes between leisure and household activities. Furthermore, it asks for occupational activities and locomotion. For further processing, the sum of leisure activity induced energy metabolism [Metabolic Equivalent of Task-hours (METH)] from age 20–80 years was calculated. Adequate validity and good reliability have already been proved in a cohort of older adults (Vogel et al., 2020).

Sleep parameters were queried by means of the Pittsburgh Sleep Quality Index (PSQI), which exhibits acceptable psychometric properties in oldest-old populations (Zhang et al., 2020). The questionnaire gathers data on seven components contributing to: (1) subjective sleep quality, (2) sleep latency, (3) sleep duration, (4) sleep efficiency, (5) sleep disturbances, (6) sleeping-pill intake, and (7) daytime sleepiness (Buysse et al., 1989). For further processing, raw data on sleep duration were utilized (hours).

Both questionnaires were completed as an interview.

Current activity levels were recorded by uniaxial accelerometers (MyWellness Key, Technogym, Gambettola, IT) worn near the hip for 7 consecutive days. Datasets of at least four valid days (wear time of at least 10 h) were considered valid (Gabrys et al., 2015). Subject to measurement was physical activity [measured as energy expended for acceleration of the body (METH)] in total and in different intensity ranges, predefined by the accelerometers (1.8–2.9 MET; 3–5.9 MET; >6 MET). The accelerometers exhibit sufficient measurement properties in cohorts of older adults *via* cross-validation (Colbert et al., 2011; Gardiner et al., 2011; Herrmann et al., 2011; Bergamin et al., 2012; Sieverdes et al., 2013).

Data Processing

Each group was further sub-stratified using cutoffs for sleep duration. Here, a subgroup who reported sleeping in recommended extents (REs) and one reporting divergent extents (DEs) was stratified in each, the IG and CG. An "optimal" sleep duration appears to depend on age and activity level of the investigated cohort as well as on the investigated outcome. Hence, we conducted a preliminary study and screened the literature for appropriate studies to determine tailored cutoffs. REs were defined as 5–7 h of sleep, as this period is found to be an intersection of results and findings of our preliminary study regarding health-beneficial sleep durations in the literature (Ferrara and de Gennaro, 2001; Gangwisch et al., 2008; Bellavia et al., 2014; Hall et al., 2015; Devore et al., 2016; Silva et al., 2016; Vogel et al., 2021b). Consequently, DEs concern any sleep duration of <5 or more than 7 h.

The LLPAQ data were translated to METH and overall activity from the age span 20–80 years was summed. Activity levels below the age of 20 were not taken into account due to an expectedly high reporting bias (Vogel et al., 2020).

The change scores of primary and secondary outcomes were computed by subtracting initial measurement values from follow-up measurement values. According to a $\sqrt{v}$-dependency shown in literature, step-length, and cadence (captured in gait analysis) were normalized to speed for further processing (Winter, 2009).

Statistics

Statistical programs used were statistical package for social sciences (SPSS) for Windows (Version 22, IBM, SPSS Inc., Chicago, IL, USA) and biometrical analyses of samples (BiAS) for Windows (Version 9.05, Goethe-University Frankfurt, Germany). A p-value of 5% was considered as a relevant cutoff for all significance testing.

Normal distribution (Shapiro-Wilk test; $p > 0.05$) and homoscedasticity (Levene test; $p > 0.05$) were both checked and given.

Baseline sub-cohort differences were analyzed by means of univariate ANOVA. Differences between genders (male vs. female), body mass index (BMI)–(18.5–24.9 vs. <18.5 & >24.9), and mild cognitive impairment (MCI)-ratings (<23 MoCA points vs. ≥23 MoCA points) were analyzed using the t-tests.

Intervention effects were evaluated using univariate ANCOVAs (4 × 1 ANCOVA) for each outcome in an intention-to-treat (ITT) analysis. Missing values were complemented by mean substitution. Randomized groups and sub-cohorts of different sleep duration were used for grouping (independent variable), and change scores of the effect estimating variables as the dependent variables. The potential effect modifiers (baseline values, habitual physical activity, and movement biography) were used as covariates.

post-hoc analyses used Fisher's Least Significant Difference (LSD) test to explore differences between individual groups.

RESULTS

Out of the 52 recruited participants, three participants were excluded for not meeting the inclusion criteria of a MoCA threshold of 17 points. Of the remaining 49 [36 women; 82.9 ± 4.5 years of age (M ± SD)] randomly assigned participants, 25 received treatment. Overall dropout from both groups comprised 12 participants (7 IG; 5 CG). Reported reasons for dropout were a lack of interest or motivation, and health status of spouses. No adverse events occurred. There was no significant difference regarding baseline measures between participants that dropped out or stuck to the intervention. Participants flow is given in **Figure 1**.

Baseline demographic characteristics are charted in **Table 2**.

The sub-cohorts differed regarding body weight and BMI at baseline [body weight: $F_{(3,48)} = 3.325$; $p = 0.039$; BMI: $F_{(3,48)} = 3.402$; $p = 0.037$]. Baseline measurement of self-report health status yielded a physical score of (mean) 42.9 ± (SD) 11.0 points and a mental score of 53.9 ± 8.3 points. The baseline cognitive capacity was 24.8 ± 2.9 points (MoCA) and fear of falls 9.9 ± 3.5 points (FES-I). Balance and leg strength testing resulted in 36.0 ± 9.5 cm at "functional-reach" and 15.7 ± 5.3 s in the "Chair-Rising" -test. Gait analyses revealed a step-length of 55.9 ±7.3 cm and cadence of 110.0 ± 17.2 steps/min, both normalized to individual gait speed. Average walking speed was 1.0 ± 0.2 m/s, double support phase was 34.1 ± 7.6%, and the walk-ratio was 5.3 ± 1.3 mm/(step/min). Baseline and postintervention sub-cohort values for each measurement are shown in **Table 3**.

The change scores of effect estimators are shown in the **Figures 2–4**. Significant differences between groups were found for change scores of cognitive capacities [$F_{(3,48)} = 5.498$, $p = 0.003$, $\eta p^2 = 0.287$]. Balance, leg strength, and gait parameters showed no sub-cohort differences (each $p > 0.05$). Changes in fear of falls (FES-I) and balance change scores differed between groups, a contribution of the covariate "current activity levels (accelerometers)" was given [fear of falls: $F_{(1,48)} = 12.961$, $p = 0.001$, $\eta p^2 = 0.240$; functional reach: $F_{(1,48)} = 4.521$, $p = 0.040$, $\eta p^2 = 0.099$]. Change scores of gait cadence differed between groups (movement biography as the significant covariate: $F_{(1,48)} = 4.545$; $p = 0.039$, $\eta p^2 = 0.100$). No other between-group-differences in any gait characteristic occurred ($p > 0.05$) (**Table 4**).

The *post-hoc* results are (as an addition to the values) illustrated in **Figures 2–4**.

DISCUSSIONS

We found that, in older adults, a multimodal exercise intervention improved cognitive capacity in the IG sleeping at REs, but not in the IG exhibiting divergent sleep durations. In the CG exhibiting divergent sleep durations, the cognitive capacity decreased. Since there were no differences between particular intervention or sleep duration groups regarding functional outcomes, both hypotheses can only be verified for cognitive capacity.

Sample Representativeness

Mean values above reference values were found for double support phase, mental health, functional-reach, and the chair-rising-test of the study cohort (+22.2% in comparison to 27.9% double support phase of the reference values; +44.0% in comparison to a summary score of 37.5 points of the reference values for the SF12 mental health score; +6.2% in comparison to the 33.9 cm seated functional reach of the reference values; +41.4% in comparison to 11.1 s of the reference values for the chair-rising-test) (Jakobsson, 2007; Thompson and Medley, 2007; Delbaere et al., 2010; Hollman et al., 2011; Bohannon et al., 2017; Carson et al., 2018; Freitas et al., 2018; Kawai et al., 2019; Rosa et al., 2019). Below-reference mean values were found for physical health (−18.7% in comparison to a summary score of 50.3 points of the reference values for the SF12 physical health score), gait speed (−20.6% in comparison to 1.26 m/s for the reference value), step-length (−8.8% in comparison to 61.3 cm step-length of the reference values), step cadence (−11.9% in comparison to 124.3 steps/min for the reference values), and walk-ratio [−11.7 in comparison to 6.0 mm/(step/min) for the reference values] (Hollman et al., 2011; Makizako et al., 2017; Bogen et al., 2018; Bergland and Strand, 2019; Kawai et al., 2019; Ramírez-Vélez et al., 2020). Conclusively, our sample is considered representative of the underlying population since the captured parameters are not consistently above or below average.

Intervention Effects and the Role of Sleep Duration

We assume the BMI differences too low and BMI values too close to healthy values to affect outcomes and to refer to BMI as an indicator of unhealthy lifestyles in the present cohort. Furthermore, the sub-cohorts exhibited common BMI scores on average with no extraordinary values in individual participants (Peralta et al., 2018).

Cognitive capacity was improved by the intervention, in particular, in the sub-cohort who slept the recommended amount. In combination with exercise, medium sleep durations appear to beneficially affect cognitive capacity. Exceptional sleep durations, on the other hand, presumably have adverse effects (Lo et al., 2016). The latter is true, irrespective of the group allocation. Seniors with medium sleep duration are, thus, trainable by multimodal exercise interventions in view of a positive influence on cognitive performance. We also found an intervention effect on cognition in participants exhibiting excessively long or comparable short sleep durations. Yet, the isolated effect of recommended sleep durations on cognitive decline may be larger

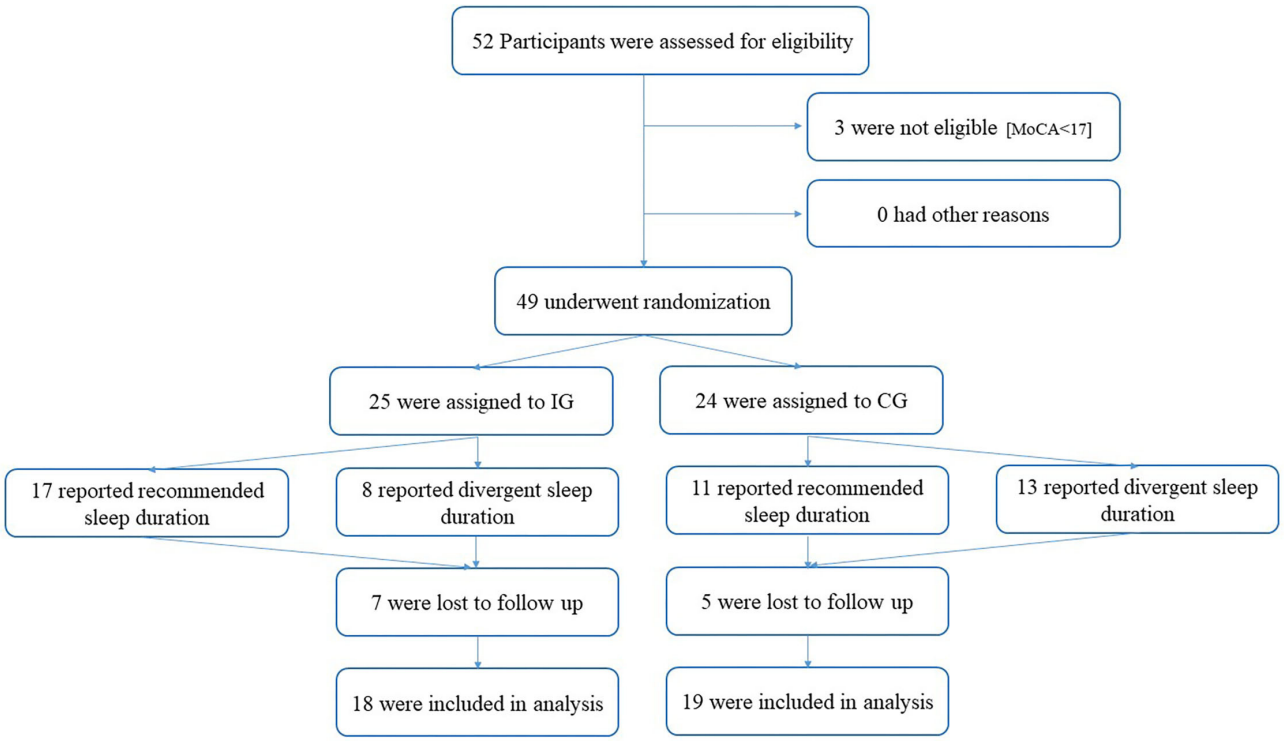


FIGURE 1 | Participant flow. IG, Intervention Group; CG, Control Group; MoCA, Montreal Cognitive Assessment).

TABLE 2 | Baseline demographic characteristics.

	Intervention group-recommended extents		Intervention group-divergent extents		Control group-recommended extents		Control group-divergent extents		Group differences	Total	
N	17		8		11		13			49	
Female	11		7		8		10		$F = 0.563$; $p = 0.645$	36	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**		**Mean**	**SD**
Age [years]	83.4	4.4	83.5	4.9	82.2	4.8	82.4	4.8	$F_{(3,48)} = 0.219$; $p = 0.882$	82.9	4.6
Height [cm]	168.8	8.2	163.9	8.2	169.1	10.5	167.5	8.2	$F_{(3,48)} = 0.689$; $p = 0.568$	167.7	8.7
Weight [Kg]	66.0	9.9	56.1	9.8	66.4	13.4	72.2	12.7	$F_{(3,48)} = 3.325$; $p = 0.039$	66.0	12.3
BMI	23.1	2.6	20.9	3.2	23.1	2.9	25.5	3.1	$F_{(3,48)} = 3.402$; $p = 0.037$	23.3	3.2
MoCA [POINTS]	25.2	3.1	24.9	2.9	25.0	2.7	24.2	2.9	$F_{(3,48)} = 0.314$; $p = 0.082$	24.8	2.9
Current activity [moves/day]	501.7	259.9	357.8	202.7	411.7	279.6	268.8	158.8	$F_{(3,48)} = 1.150$; $p = 0.424$	369.9	233.7
Lifetime-activity [Total METh]	189,371	176,874	127,556	115,693	170,187	163,531	301,163	276,809	$F_{(3,48)} = 0.461$; $p = 0.717$	199,707	188,559

CM, Centimeters; KG, Kilogramm; MOCA, Montreal Cognitive Assessment; METH, Metabolic Equivalent-Hours; SD, Standard Deviation.

than the isolated effect of solely participating in the exercise intervention. Change scores of the groups exhibiting both or no factors indicate combinable effects of healthy sleep durations and exercise interventions.

The CG sub-cohort sleeping <5 h or more than 7 h exhibited the largest decline of the MoCA score in the 16-week study period. The MoCA scores of three participants from this sub-cohort showed baseline values above and postintervention values below a MCI cutoff score of 23 points (Thomann et al., 2020). Contrary, the IG sleeping at REs shows gains in cognitive capacity, contrasting the usual age-related decline. Two participants of this group exhibited baseline values below and postintervention values above the MCI cutoff (23 points) (Thomann et al., 2020). Both of these groups exhibit change scores above the minimum detectable change for the MoCA (Feeney et al., 2016). According to this, the combination

TABLE 3 | Effect estimator values at baseline and post-intervention for each sub-cohort.

	Intervention group-recommended extents		Intervention group-divergent extents		Control group-recommended extents		Control group-divergent extents	
Time of measurement	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**
Montreal Cognitive Assessment [points]	25.2 ± 3.1	26.4 ± 3.1	24.9 ± 2.9	23.6 ± 4.2	25.0 ± 2.7	24.7 ± 2.8	24.2 ± 2.9	21.0 ± 3.9
Falls efficacy scale international [points]	9.7 ± 2.5	9.5 ±1.9	10.0 ± 3.5	9.4 ± 2.5	9.7 ± 3.7	9.3 ± 3.8	10.4 ± 4.8	10.9 ± 4.9
Shortform 12 mental score [points]	54.5 ± 7.9	53.9 ± 8.2	52.9 ± 9.8	54.6 ± 11.0	57.7 ± 4.8	55.1 ± 7.9	51.6 ± 10.0	53.0 ± 11.3
Shortform 12 physical score [points]***	46.3 ± 10.7	46.7 ± 10.7	43.3 ± 9.6	42.3 ± 15.3	45.3 ± 10.1	48.8 ± 12.4	35.2 ± 10.3	36.3 ± 10.7
Funtional reach test [cm]	37.5 ± 8.5	38.5 ± 7.7	33.2 ± 12.6	36.6 ± 8.9	38.4 ± 6.7	36.4 ± 5.3	33.2 ± 10.5	33.8 ± 13.2
Chair-rise test [seconds]	15.8 ± 7.5	15.6 ± 8.1	16.8 ± 6.5	14.4 ± 5.3	15.0 ± 3.6	12.0 ± 1.7	15.0 ± 4.9	12.9 ± 4.0
Gait–speed [m/s]	1.03 ± 0.21	1.12 ± 0.12	0.93 ± 0.30	1.07 ± 0.21	1.01 ± 0.26	0.98 ± 0.33	0.94 ± 0.23	0.98 ± 0.27
Gait–cadence [steps/min]	103 ± 8	103 ± 7	110 ± 18	108 ± 12	107 ± 11	107 ± 17	122 ± 27	125 ± 25
Gait–step-length [cm]	58.7 ± 4.4	58.6 ± 4.2	55.8 ± 8.8	55.9 ± 5.5	56.5 ± 6.0	57.2 ± 7.8	51.1 ± 10.1	49.4 ± 8.9
Gait–double support phase [% of gait cycle]	32.5 ± 4.2	32.1 ± 4.1	34.6 ± 7.0	34.2 ± 7.7	31.9 ± 7.1	32.4 ± 5.7	39.5 ± 10.7	38.2 ± 10.4

****Sub-cohort differences at baseline measurement ($p < 0.05$).*
CM, Centimeters; M/S, Meters per second; Steps/Min, Steps per Minute.

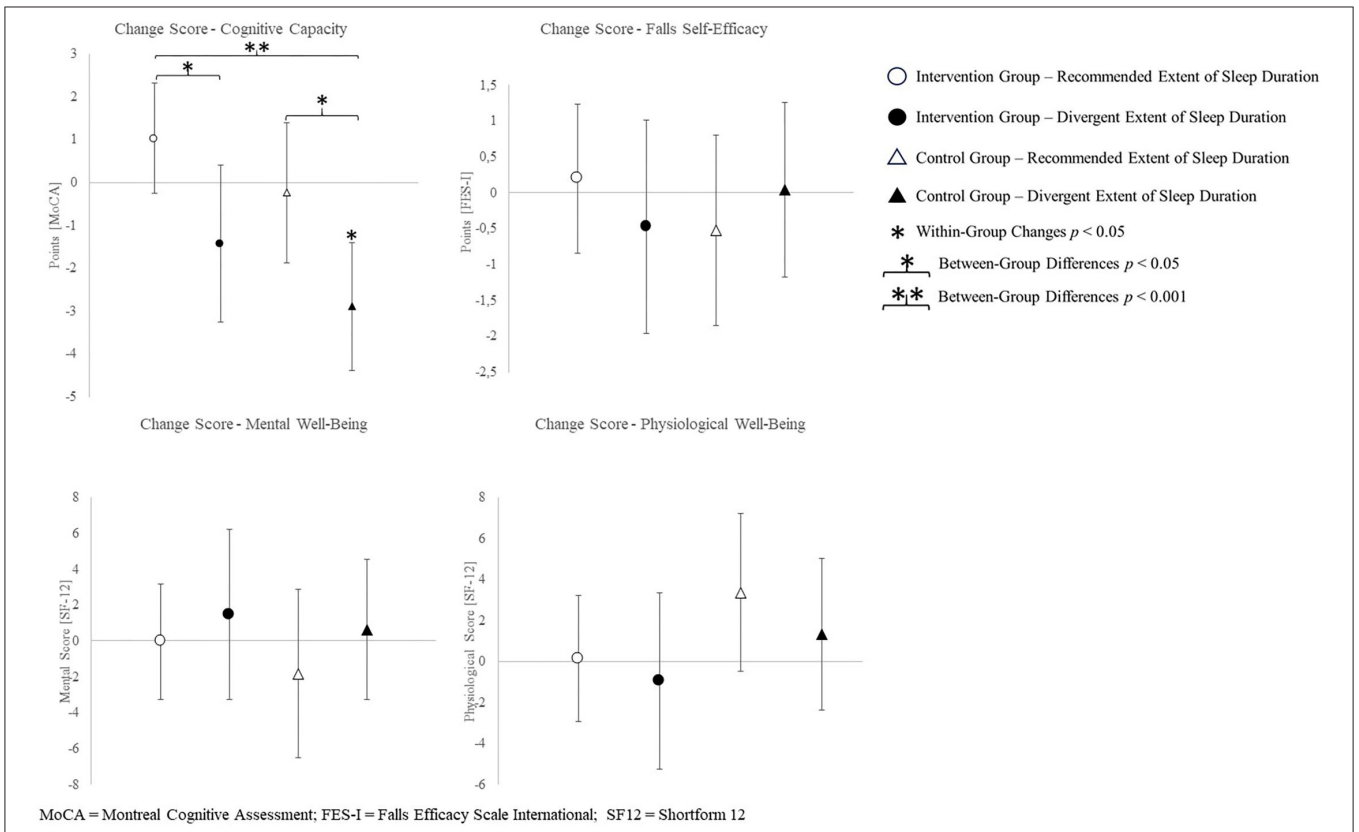


FIGURE 2 | Change scores of the primary outcome cognitive capacity, measured by the Montreal cognitive assessment and secondary outcomes fear of falling and health-related quality of life.

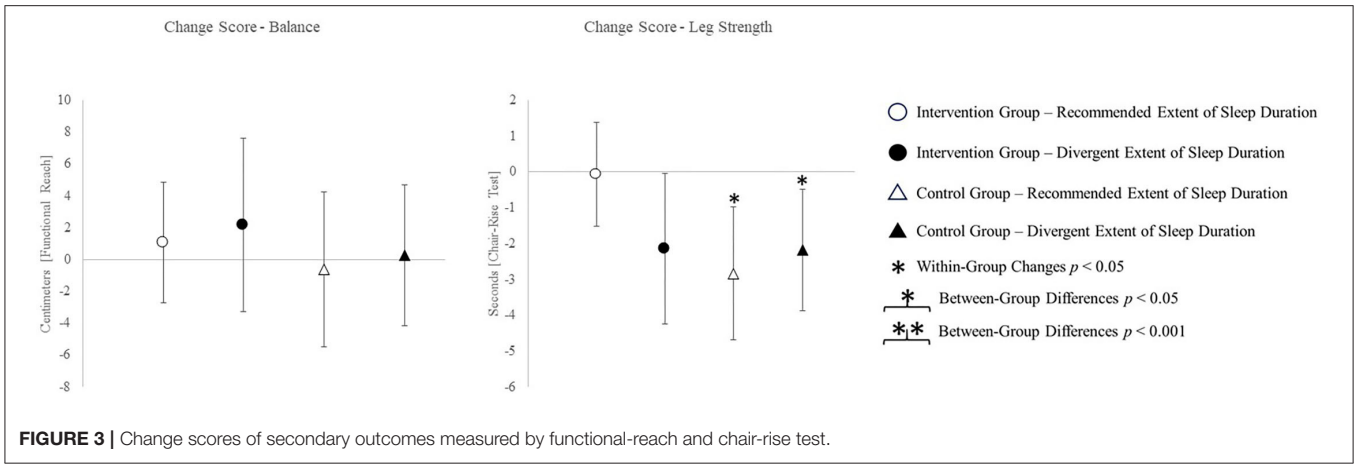


FIGURE 3 | Change scores of secondary outcomes measured by functional-reach and chair-rise test.

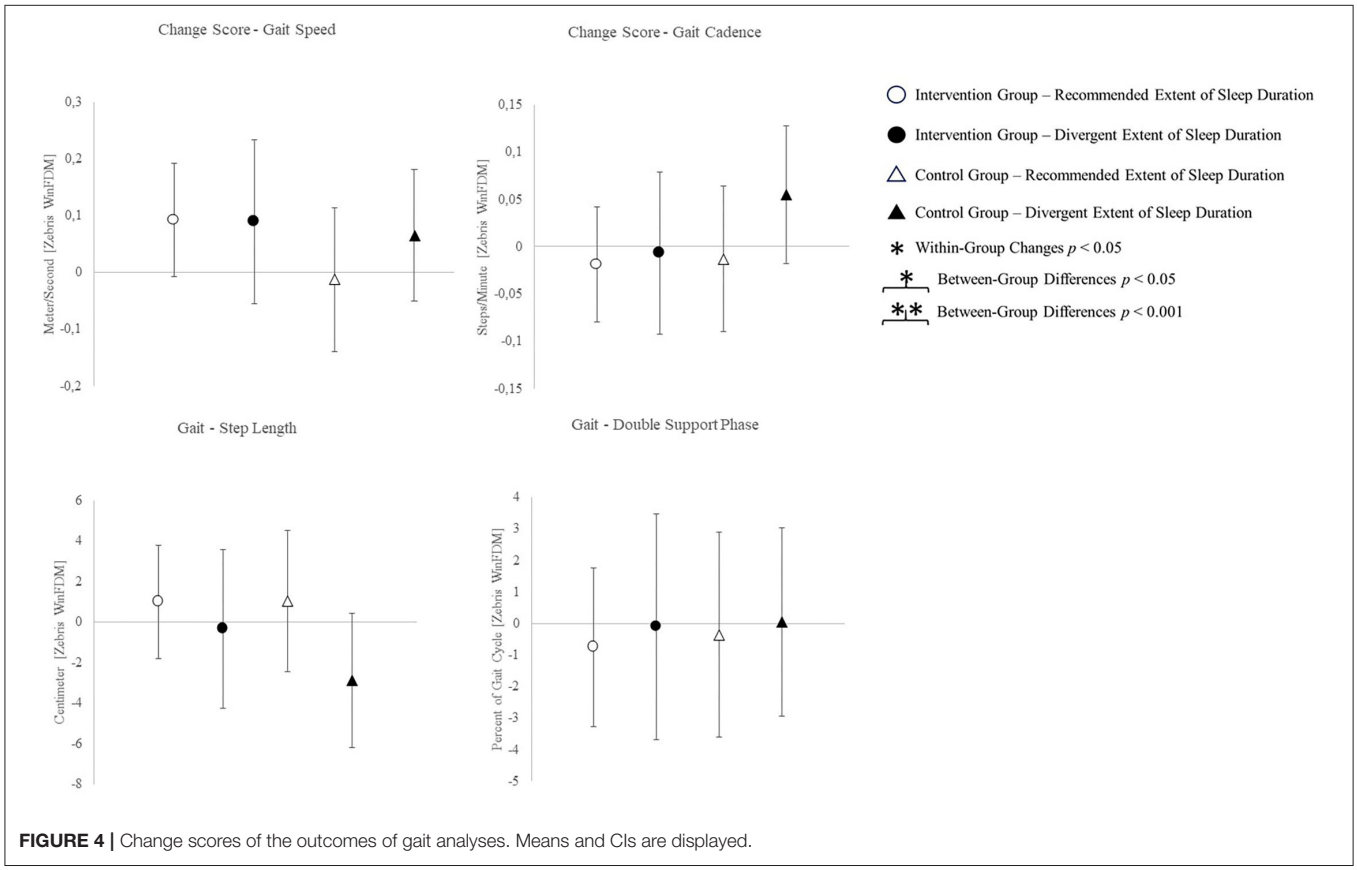


FIGURE 4 | Change scores of the outcomes of gait analyses. Means and CIs are displayed.

of a healthy sleep duration and participation in an exercise intervention yields improvements regarding cognitive capacity in the present cohort.

Incongruent results regarding exercise induced increases in self-reported health scores have been reported (Dekker-van Weering et al., 2017). The present results align with findings on exercise interventions exclusively affecting extraordinary low levels of self-reported well-being (Moriyama et al., 2020). A predictive value of sleep complaints on self-reported health and an inverse association of sleep duration and physical health scores is indicated (Reid et al., 2006; Johnson et al., 2017). Sleep efficiency on the other hand shows a positive association with both, the physical and mental aspects of self-reported health (Reid et al., 2006). Therefore, consistent sleep values used for group allocation possibly account for low mental health changes. An intervention targeting changes in sleep quality, however, might lead to changes in sleep parameters and mental health likewise. Health-related quality of life physical sum score, the functional reach ability, and most gait characteristics showed no between-group or intervention effects. Exclusively, the CG subcohort sleeping <5 h or more than 7 h shows a reduction in step-length and increase in cadence. The walk-ratio, calculated

from step-length and cadence (step-length/cadence=walk-ratio), has been reported to relate to the risk of falls in older adults (Nakakubo et al., 2018). Besides multicomponent interventions, certain sleep durations appear to affect the risk of falls as well (Fu et al., 2019). However, the subgroup neither participating in the intervention, nor sleeping at REs, exhibits an increased walk-ratio, which indicates a decreased risk of falls (walk-ratio change scores: IG-RE −0.369 ± 7.152; IG-DE −0.426 ± 11.807; CG-RE 1.830 ± 16.954; CG-DE −3.101 ± 14.600).

The missing relationship between concerns of falls and sleep parameters in the present study does not align with the findings of related studies. Divergent measurement methods might cause the lack of influence by sleep duration. Whereas, the present study measured the concerns of falling by questionnaire, related studies measured the risk of falling by means of objective measures of balance performance (Hita-Contreras et al., 2018; Serrano-Checa et al., 2020). Hence, subjective concerns and the objective risk of falling might differ regarding their relationship toward sleep parameters. Intervention efficiency on FES-I change scores might be owed to the cognitive status of participants. An even more detailed look into the MoCA data regarding the MCI cutoff criteria revealed a different profit in terms of FES-I scores in participants above and below the MCI cutoff (23 points) (Thomann et al., 2020). Participants with a lower risk of MCI (≥23 points) improved (lowered) their FES-I score throughout the intervention period, whereas participants presenting with a higher risk of MCI (<23 points) exhibited worsened (increased) FES-I scores.

Health-related quality of life—mental health showed no changes in the subgroup participating in the intervention and sleeping at REs, or in the CG sleeping excessively short or long. A tendential increase in the IG sleeping excessively short or long and a decrease in the CG sub-cohort who sleeps the recommended amount is indicated. This is somewhat surprising. Regarding the present data, the multicomponent intervention might enhance mental health, whereas sleeping at recommended durations in the present study might be related to adverse effects. Hence, the effects of the intervention and recommended sleep durations seem to cancel each other out in the present results, as well as divergent sleep durations and being allocated to the CG does. Related studies report poorer mental health outcomes in association with a reduction in sleep duration (Tang et al., 2017). As limits of sleep duration differ between studies (<6 h, 6–8 h, >8 h in related studies; <5 h and >7 h, 5–7 h in the present study), the "reduction" in sleep duration might be equivalent to the recommended sleep durations in the present study (Tang et al., 2017). Consequently, considering the variability of "optimal" sleep durations, the allegedly differing results might align nevertheless. Especially, regarding the outcomes on mental health, an active CG might have been useful. An active CG might have helped discriminating the effect of social interaction within the group and with the trainer from the effects evoked by the intervention itself.

Even more surprising was the improvement in the chair-rise performance in each group except the IG with the recommended amount of sleep. The sleep duration being most helpful for physical strength gains appears to differ from the range we

TABLE 4 | ANCOVA results regarding intervention effects for the independent variable (group allocation) and covariates (baseline values, movement biography, and current activity).

	Group allocation			Baseline values			Movement biography			Current activity			Estimated marginal means			
	F	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*M*	*SE*	*UL*	*LL*
Montreal cognitive assessment	5.498	0.003	0.287	4.320	0.044	0.095	3.376	0.073	0.076	2.131	0.152	0.049	−0.884	0.383	−1.657	−0.112
Falls efficacy scale international	0.357	0.785	0.025	15.780	<0.001	0.278	0.495	0.486	0.012	12.961	0.001	0.240	−0.189	0.312	−0.819	0.441
Shortform 12 mental score	0.403	0.752	0.029	2.160	0.149	0.050	0.622	0.435	0.015	2.000	0.165	0.047	0.051	0.994	−1.956	2.058
Shortform 12 physical score	0.910	0.445	0.062	0.065	0.801	0.002	0.178	0.675	0.004	0.948	0.336	0.023	0.957	0.901	−0.862	2.777
Funtional reach test	0.220	0.882	0.016	16.466	<0.001	0.287	0.272	0.605	0.007	4.521	0.040	0.099	0.710	1.136	−1.584	3.004
Chair–rise test	2.307	0.091	0.144	5.458	0.024	0.117	0.569	0.455	0.014	2.724	0.106	0.062	−1.813	0.437	−2.695	−0.930
Gait–speed	0.642	0.593	0.045	24.803	<0.001	0.377	3.747	0.060	0.084	3.162	0.083	0.072	0.058	0.030	−0.002	0.119
Gait–cadence	0.865	0.467	0.060	11.706	<0.001	0.222	4.545	0.039	0.100	0.149	0.702	0.004	0.004	0.018	−0.032	0.040
Gait–step-length	1.221	0.314	0.082	18.463	<0.001	0.311	4.517	0.040	0.099	0.157	0.694	0.004	−0.298	0.819	−1.952	1.355
Gait–double support phase	0.062	0.980	0.005	9.413	0.004	0.187	0.955	0.334	0.023	0.495	0.485	0.012	−0.295	0.750	−1.809	1.219

M, Mean; SE, Standard Error; UL, Upper Limit 95% Confidence Interval; LL, Lower Limit 95% Confidence Interval.

defined as recommended. Related studies report lower physical performance in relation to excessively long sleep durations (Fu et al., 2017). However, the sleep durations utilized for grouping do not even overlap between the present and related study (7–9 h in the related study; 5–7 h in the present study). Therefore, inconclusive results regarding leg strength might be owed to divergent "optimal" sleep durations between different outcomes.

Modifying Factors

Habitual physical activity and movement history were found to be relevant covariates when exercise effects in different sleep-duration cohorts are investigated. This potential interaction complies with the results of related studies (Newton, 2001; Seco et al., 2013; Pau et al., 2014; Dawe et al., 2018). In participants exhibiting high lifetime but low current activity levels, former benefits in balance skills might have already deteriorated (Rosa et al., 2019). The present cohort predominantly moved at low intensities (89.6% of the captured activity), complying with literature on activity of older adults (Troiano et al., 2008). The lack of moderate to vigorous physical activity (MVPA) prevents analyses of their isolated impact on the investigated outcomes, which might provide beneficial insights.

While most studies use a self-report measure for physical activity or objective measures for isolated MVPA, we objectively captured total amounts of physical activity. Therefore, subliminal activity not reported in questionnaires might confound findings on gait parameter relationships (Ciprandi et al., 2017; Hamacher et al., 2018).

Limitations and Future Studies

To assess balance, we utilized the seated functional-reach test in favor of fall risk minimization. As the seated version is less demanding than the standing version, it might have caused a ceiling effect in balance performance. Even though the selection of exercises, intensity, and tempo was tailored to our cohort, individual and continuous customization of the exercise load possibly enhances intervention effects in cohorts of heterogenous fitness levels (Herold et al., 2019). Since a highly personalized training intensity does not reflect common practice, an individual approach might facilitate more definitive results at the expense of generalizability and practical applicability. Although sleep parameter improvements are rather evoked by moderate than high-intensity exercise, there might be a reverse relationship to other entities (Bullock et al., 2020). We recommend further studies to integrate complementing objective sleep measurements and individual, subjective reports on perceived intervention intensity for continuous adjustment. The number of participants dropping out exceeded the number of additionally recruited participants in our cohort. For longitudinal studies investigating older adults lasting 4 months or more, it might be advisable to recruit more than the recommended additional 20%. Furthermore, data on sustainability acquired from a follow-up measurement, as well as on autonomous continuation of comparable activities, might provide further insight regarding the effectiveness of interventions in terms of health promotion. An additional active CG might be useful for further research in terms of discriminating the effect of social interaction within the group and with the trainer from the effects evoked by the intervention itself.

Generalizability and Practical Relevance

Since balance, leg strength, and health status of the present cohort corresponds to age-matched general population, we consider our results to be generalizable. As high degrees of personalization do not comply with usual group-based health promotion offerings, they are not suitable for increasing generalizability. Some imbalance between performance status and intervention-induced physical demands of individuals is not uncommon and favors generalizability (Uemura et al., 2013). Our findings on covariates appear to be generalizable as all measurements of utilized variables represent usual dimensions.

CONCLUSIONS

We found health beneficial effects of a multicomponent intervention in consideration of habitual sleep duration, habitual physical activity, and lifetime activity. In particular, sleep duration and multimodal exercises affect cognitive capacity synergistically. Fear of falls and balance change scores depend on current physical activity levels, whereas gait cadence changes depend on lifetime physical activity. From a general perspective, multivariate analyses of intervention effects are warranted to systematically assess interactions as well as mediating and moderating effects of multiple health contributing factors within one analytic approach. Research on synergistic, moderating, and mediating effects of health factors possibly enables insight into their mechanisms of action (Ohrnberger et al., 2017; Whibley et al., 2019). Such findings might enable a deeper understanding of differing effectiveness between interventions and help to enhance existing interventions (Janssen et al., 2020). Furthermore, besides physical activity guidelines, promotion of sleep behavior recommendations in older adults might benefit the maintenance of a healthy lifestyle and in particular, contribute to the prevention of cognitive decline.

DATA AVAILABILITY STATEMENT

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Goethe-University, Department of Psychology and Sports Sciences, Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

LV, OV, and DN: study concept and design, analysis and interpretation of data, and preparation of manuscript. OV: acquisition of data. All authors have read and approved the manuscript.

REFERENCES

Arora, T., Oates, A., Lynd, K., and Musselman, K. E. (2020). Current state of balance assessment during transferring, sitting, standing and walking activities for the spinal cord injured population: A systematic review. *J. Spinal Cord Med.* 43, 10–23. doi: 10.1080/10790268.2018.1481692

Bellavia, A., Åkerstedt, T., Bottai M, Wolk, A., and Orsini, N. (2014). Sleep duration and survival percentiles across categories of physical activity. *Am. J. Epidemiol.* 179, 484–491. doi: 10.1093/aje/kwt280

Bergamin, M., Ermolao, A., Sieverdes, J. C., Zaccaria, M., and Zanuso, S. (2012). Validation of the mywellness key in walking and running speeds. *J. Sports Sci. Med*. 11, 57–63.

Bergland, A., and Strand, B. H. (2019). Norwegian reference values for the Short Physical Performance Battery (SPPB): the Tromsø Study. *BMC Geriatr*. 19:216. doi: 10.1186/s12877-019-1234-8

Bogen, B., Moe-Nilssen, R., Ranhoff, A. H., and Aaslund, M. K. (2018). The walk ratio: Investigation of invariance across walking conditions and gender in community-dwelling older people. *Gait Posture* 61, 479–482. doi: 10.1016/j.gaitpost.2018.02.019

Bohannon, R. W., Wolfson, L. I., and White, W. B. (2017). Functional reach of older adults: normative reference values based on new and published data. *Physiotherapy* 103, 387–391. doi: 10.1016/j.physio.2017.03.006

Bullock, A., Kovacevic, A., Kuhn, T., and Heisz, J. J. (2020). Optimizing sleep in older adults: where does high-intensity interval training fit? *Front. Psychol.* 11:576316. doi: 10.3389/fpsyg.2020.576316

Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., and Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Res.* 28, 193–213. doi: 10.1016/0165-1781(89)90047-4

Cadore, E. L., Rodríguez-Mañas, L., Sinclair, A., and Izquierdo, M. (2013). Effects of different exercise interventions on risk of falls, gait ability, and balance in physically frail older adults: a systematic review. *Rejuvenation Res.* 16, 105–114. doi: 10.1089/rej.2012.1397

Carson, N., Leach, L., and Murphy, K. J. (2018). A re-examination of Montreal Cognitive Assessment (MoCA) cutoff scores. *Int. J. Geriatr. Psychiatry*. 33, 379–388. doi: 10.1002/gps.4756

Chaput, J.-P., Dutil, C., Featherstone, R., Ross, R., Giangregorio, L., Saunders, T. J., et al. (2020). Sleep timing, sleep consistency, and health in adults: a systematic review. *Appl. Physiol. Nutr. Metab*. 45, S232–S247. doi: 10.1139/apnm-2020-0032

Chatfield, M. D., Brayne, C. E., and Matthews, F. E. (2005). A systematic literature review of attrition between waves in longitudinal studies in the elderly shows a consistent pattern of dropout between differing studies. *J. Clin. Epidemiol.* 58, 13–19. doi: 10.1016/j.jclinepi.2004.05.006

Ciprandi, D., Bertozzi, F., Zago, M., Ferreira, C. L. P., Boari, G., Sforza, C., et al. (2017). Study of the association between gait variability and physical activity. *Eur. Rev. Aging Phys. Act*. 14:19. doi: 10.1186/s11556-017-0188-0

Colbert, L. H., Matthews, C. E., Havighurst, T. C., Kim, K., and Schoeller, D. A. (2011). Comparative validity of physical activity measures in older adults. *Med. Sci. Sports Exerc*. 43, 867–876. doi: 10.1249/MSS.0b013e3181fc7162

Cordes, T., Bischoff, L. L., Schoene, D., Schott, N., Voelcker-Rehage, C., Meixner, C., et al. (2019). A multicomponent exercise intervention to improve physical functioning, cognition and psychosocial well-being in elderly nursing home residents: a study protocol of a randomized controlled trial in the PROCARE (prevention and occupational health in long-term care) project. *BMC Geriatr*. 19:369. doi: 10.1186/s12877-019-1386-6

Daskalopoulou, C., Stubbs, B., Kralj, C., Koukounari, A., Prince, M., and Prina, A. M. (2017). Physical activity and healthy ageing: A systematic review and meta-analysis of longitudinal cohort studies. *Ageing Res. Rev*. 38, 6–17. doi: 10.1016/j.arr.2017.06.003

Dawe, R. J., Leurgans, S. E., Yang, J., Bennett, J. M., Hausdorff, J. M., Lim, A. S., et al. (2018). Association between quantitative gait and balance measures and total daily physical activity in community-dwelling older adults. *J. Gerontol. A Biol. Sci. Med. Sci*. 73, 636–642. doi: 10.1093/gerona/glx167

Dekker-van Weering, M., Jansen-Kosterink, S., Frazer, S., and Vollenbroek-Hutten, M. (2017). User experience, actual use, and effectiveness of an information communication technology-supported home exercise program for pre-frail older adults. *Front. Med.* 4:208. doi: 10.3389/fmed.2017.00208

Delbaere, K., Close, J. C. T., Mikolaizak, A. S., Sachdev, P. S., Brodaty, H., and Lord, S. R. (2010). The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study. *Age Ageing* 39, 210–216. doi: 10.1093/ageing/afp225

Devore, E. E., Grodstein, F., and Schernhammer, E. S. (2016). Sleep duration in relation to cognitive function among older adults: a systematic review of observational studies. *Neuroepidemiology* 46, 57–78. doi: 10.1159/000442418

Dolezal, B. A., Neufeld, E. V., Boland, D. M., Martin, J. L., and Cooper, C. B. (2017). Interrelationship between sleep and exercise: a systematic review. *Adv. Prev. Med*. 2017:1364387. doi: 10.1155/2017/1364387

Drixler, K., Morfeld, M., Glaesmer, H., Brähler, E., and Wirtz, M. A. (2020). Validierung der Messung gesundheitsbezogener Lebensqualität mittels des Short-Form-Health-Survey-12 (SF-12 Version 2.0) in einer deutschen Normstichprobe. *Z. Psychosom. Med. Psychother*. 66, 272–286. doi: 10.13109/zptm.2020.66.3.272

Feeney, J., Savva, G. M., O'Regan, C., King-Kallimanis, B., Cronin, H., and Kenny, R. A. (2016). Measurement error, reliability, and minimum detectable change in the mini-mental state examination, montreal cognitive assessment, and color trails test among community living middle-aged and older adults. *J. Alzheimers. Dis*. 53, 1107–1114. doi: 10.3233/JAD-160248

Ferrara, M., and de Gennaro, L. (2001). How much sleep do we need? *Sleep Med. Rev*. 5, 155–179. doi: 10.1053/smrv.2000.0138

Fessel, M. M., Mann, M., Miyawaki, C. E., and Rosenberg, D. E. (2017). Multicomponent interventions and cognitive health: a scoping review. *J. Gerontol. Nurs*. 43, 39–48. doi: 10.3928/00989134-20170131-01

Freiberger, E., Blank, W. A., Salb, J., Geilhof, B., Hentschke, C., Landendoerfer, P., et al. (2013). Effects of a complex intervention on fall risk in the general practitioner setting: a cluster randomized controlled trial. *Clin. Interv. Aging*. 8, 1079–1088. doi: 10.2147/CIA.S46218

Freitas, S., Batista, S., Afonso, A. C., Simões, M. R., Sousa, L., de, Cunha, L., et al. (2018). The Montreal Cognitive Assessment (MoCA) as a screening test for cognitive dysfunction in multiple sclerosis. *Appl. Neuropsychol. Adult*. 25, 57–70. doi: 10.1080/23279095.2016.1243108

Friedman, S. M., Mulhausen, P., Cleveland, M. L., Coll, P. P., Daniel, K. M., Hayward, A. D., et al. (2019). healthy aging: american geriatrics society white paper executive summary. *J. Am. Geriatr. Soc*. 67, 17–20. doi: 10.1111/jgs.15644

Fu, L., Jia, L., Zhang, W., Han, P., Kang, L., Ma, Y., et al. (2017). The association between sleep duration and physical performance in Chinese community-dwelling elderly. *PLoS ONE* 12:e0174832. doi: 10.1371/journal.pone.0174832

Fu, L., Yu, X., Zhang, W., Han, P., Kang, L., Ma, Y., et al. (2019). The relationship between sleep duration, falls, and muscle mass: a cohort study in an elderly Chinese population. *Rejuvenation Res*. 22, 390–398. doi: 10.1089/rej.2018.2102

Gabrys, L., Thiel, C., Tallner, A., Wilms, B., Müller, C., Kahlert, D., et al. (2015). Akzelerometrie zur Erfassung körperlicher Aktivität. *Sportwiss* 45, 1–9. doi: 10.1007/s12662-014-0349-5

Gadie, A., Shafto, M., Leng, Y., and Kievit, R. A. (2017). How are age-related differences in sleep quality associated with health outcomes? An epidemiological investigation in a UK cohort of 2406 adults. *BMJ Open* 7:e014920. doi: 10.1136/bmjopen-2016-014920

Gangwisch, J. E., Heymsfield, S. B., Boden-Albala, B., Buijs, R. M., Kreier, F., Opler, M. G., et al. (2008). Sleep duration associated with mortality in elderly, but not middle-aged, adults in a large US sample. *Sleep* 31, 1087–1096. doi: 10.5665/sleep/31.8.1087

Gardiner, P. A., Clark, B. K., Healy, G. N., Eakin, E. G., Winkler, E. A. H., and Owen, N. (2011). Measuring older adults' sedentary time: reliability, validity, and responsiveness. *Med. Sci. Sports Exerc*. 43, 2127–2133. doi: 10.1249/MSS.0b013e31821b94f7

Gheysen, F., Poppe, L., DeSmet, A., Swinnen, S., Cardon, G., Bourdeaudhuij, I., et al. (2018). Physical activity to improve cognition in older adults: can physical activity programs enriched with cognitive challenges enhance the effects? A systematic review and meta-analysis. *Int. J. Behav. Nutr. Phys. Act*. 15:63. doi: 10.1186/s12966-018-0697-x

Gopinath, B., Kifley, A., Flood, V. M., and Mitchell, P. (2018). Physical activity as a determinant of successful aging over ten years. *Sci. Rep*. 8:10522. doi: 10.1038/s41598-018-28526-3

Hall, M. H., Smagula, S. F., Boudreau, R. M., Ayonayon, H. N., Goldman, S. E., Harris, T. B., et al. (2015). Association between sleep duration and mortality is mediated by markers of inflammation and health in older adults: the Health,

Aging and Body Composition Study. *Sleep* 38, 189–195. doi: 10.5665/sleep.4394

Hamacher, D., Liebl, D., Hödl, C., Heßler, V., Kniewasser, C. K., Thönnessen, T., et al. (2018). Gait stability and its influencing factors in older adults. *Front. Physiol*. 9:1955. doi: 10.3389/fphys.2018.01955

Haskell, W. L., Blair, S. N., and Hill, J. O. (2009). Physical activity: health outcomes and importance for public health policy. *Prevent. Med*. 49, 280–282. doi: 10.1016/j.ypmed.2009.05.002

Herold, F., Müller, P., Gronwald, T., and Müller, N. G. (2019). Dose-response matters! - A perspective on the exercise prescription in exercise-cognition research. *Front. Psychol*. 10:2338. doi: 10.3389/fpsyg.2019.02338

Herrmann, S. D., Hart, T. L., Lee, C. D., and Ainsworth, B. E. (2011). Evaluation of the MyWellness Key accelerometer. *Br. J. Sports Med*. 45, 109–113. doi: 10.1136/bjsm.2009.062182

Hita-Contreras, F., Zagalaz-Anula, N., Martínez-Amat, A., Cruz-Díaz, D., Sánchez-Montesinos, I., Aibar-Almazán, A., et al. (2018). Sleep quality and its association with postural stability and fear of falling among Spanish postmenopausal women. *Menopause* 25, 62–69. doi: 10.1097/GME.0000000000000941

Hollman, J. H., McDade, E. M., and Petersen, R. C. (2011). Normative spatiotemporal gait parameters in older adults. *Gait Posture* 34, 111–118. doi: 10.1016/j.gaitpost.2011.03.024

Jadczak, A. D., Makwana, N., Luscombe-Marsh, N., Visvanathan, R., and Schultz, T. J. (2018). Effectiveness of exercise interventions on physical function in community-dwelling frail older people: an umbrella review of systematic reviews. *JBI Database System Rev. Implement Rep*. 16, 752–775. doi: 10.11124/JBISRIR-2017-003551

Jakobsson, U. (2007). Using the 12-item Short Form health survey (SF-12) to measure quality of life among older people. *Aging Clin. Exp. Res*. 19, 457–464. doi: 10.1007/BF03324731

Janssen, I., Clarke, A. E., Carson, V., Chaput, J.-P., Giangregorio, L. M., Kho, M. E., et al. (2020). A systematic review of compositional data analysis studies examining associations between sleep, sedentary behaviour, and physical activity with health outcomes in adults. *Appl. Physiol. Nutr. Metab*. 45, S248–S257. doi: 10.1139/apnm-2020-0160

Johnson, S. T., Thiel, D., Al Sayah, F., Mundt, C., Qiu, W., Buman, M. P., et al. (2017). Objectively measured sleep and health-related quality of life in older adults with type 2 diabetes: a cross-sectional study from the Alberta's Caring for Diabetes Study. *Sleep Health* 3, 102–106. doi: 10.1016/j.sleh.2016.12.002

Kawai, H., Taniguchi, Y., Seino, S., Sakurai, R., Osuka, Y., Obuchi, S., et al. (2019). Reference values of gait parameters measured with a plantar pressure platform in community-dwelling older Japanese adults. *Clin. Interv. Aging*. 14, 1265–1276. doi: 10.2147/CIA.S213216

Kempen, G. I., J. M., Todd CJ, van Haastregt, J. C. M., Zijlstra, G. A. R., Beyer, N., Freiberger, E., et al. (2007). Cross-cultural validation of the Falls Efficacy Scale International (FES-I) in older people: results from Germany, the Netherlands and the UK were satisfactory. *Disabil. Rehabil*. 29, 155–162. doi: 10.1080/09638280600747637

Kempen, G. I., J. M., Yardley L, van Haastregt, J. C. M., Zijlstra, G. A. R., Beyer, N., Hauer, K., et al. (2008). The Short FES-I: a shortened version of the falls efficacy scale-international to assess fear of falling. *Age Ageing* 37, 45–50. doi: 10.1093/ageing/afm157

Knobf, M. T., Thompson, A. S., Fennie, K., and Erdos, D. (2014). The effect of a community-based exercise intervention on symptoms and quality of life. *Cancer Nurs*. 37, E43–E50. doi: 10.1097/NCC.0b013e318288d40e

Landi, F., Calvani, R., Picca, A., Tosato, M., Martone, A. M., D'Angelo, E., et al. (2018). Impact of habitual physical activity and type of exercise on physical performance across ages in community-living people. *PLoS ONE* 13:e0191820. doi: 10.1371/journal.pone.0191820

Lo, J. C., Groeger, J. A., Cheng, G. H., Dijk, D.-J., and Chee, M. W. L. (2016). Self-reported sleep duration and cognitive performance in older adults: a systematic review and meta-analysis. *Sleep Med*. 17, 87–98. doi: 10.1016/j.sleep.2015.08.021

Makizako, H., Shimada, H., Doi, T., Tsutsumimoto, K., Nakakubo, S., Hotta, R., et al. (2017). Predictive cutoff values of the five-times sit-to-stand test and the timed "Up & Go" test for disability incidence in older people dwelling in the community. *Phys. Ther*. 97, 417–424. doi: 10.2522/ptj.20150665

McClain, J. J., Lewin, D. S., Laposky, A. D., Kahle, L., and Berrigan, D. (2014). Associations between physical activity, sedentary time, sleep duration and daytime sleepiness in US adults. *Prevent. Med*. 66, 68–73. doi: 10.1016/j.ypmed.2014.06.003

Mehmet, H., Yang, A. W. H., and Robinson, S. R. (2020). What is the optimal chair stand test protocol for older adults? A systematic review. *Disabil. Rehabil*. 42, 2828–2835. doi: 10.1080/09638288.2019.1575922

Melo, T. A., de, Duarte, A. C. M., Bezerra, T. S., França, F., Soares, N. S., and Brito, D. (2019). Teste de Sentar-Levantar Cinco Vezes: segurança e confiabilidade em pacientes idosos na alta da unidade de terapia intensiva. *Rev. Bras. Ter. Intensiva*. 31, 27–33. doi: 10.5935/0103-507X.20190006

Mendelson, M., Borowik, A., Michallet, A.-S., Perrin, C., Monneret, D., Faure, P., et al. (2016). Sleep quality, sleep duration and physical activity in obese adolescents: effects of exercise training. *Pediatr. Obes*. 11, 26–32. doi: 10.1111/ijpo.12015

Mewborn, C. M., Lindbergh, C. A., and Stephen Miller, L. (2017). Cognitive interventions for cognitively healthy, mildly impaired, and mixed samples of older adults: a systematic review and meta-analysis of randomized-controlled trials. *Neuropsychol. Rev*. 27, 403–439. doi: 10.1007/s11065-017-9350-8

Moriyama, N., Omata, J., Sato, R., Okazaki, K., and Yasumura, S. (2020). Effectiveness of group exercise intervention on subjective well-being and health-related quality of life of older residents in restoration public housing after the great East Japan earthquake: A cluster randomized controlled trial. *Int. J. Disaster Risk Reduct*. 46:101630. doi: 10.1016/j.ijdrr.2020.101630

Nakakubo, S., Doi, T., Makizako, H., Tsutsumimoto, K., Hotta, R., Kurita, S., et al. (2018). Association of walk ratio during normal gait speed and fall in community-dwelling elderly people. *Gait Posture* 66, 151–154. doi: 10.1016/j.gaitpost.2018.08.030

Newton, R. A. (2001). Validity of the multi-directional reach test: a practical measure for limits of stability in older adults. *J. Gerontol. A Biol. Sci. Med. Sci*. 56, M248–M252. doi: 10.1093/gerona/56.4.m248

Northey, J. M., Cherbuin, N., Pumpa, K. L., Smee, D. J., and Rattray, B. (2018). Exercise interventions for cognitive function in adults older than 50: a systematic review with meta-analysis. *Br. J. Sports Med*. 52, 154–160. doi: 10.1136/bjsports-2016-096587

Ohayon, M. M., Carskadon, M. A., Guilleminault, C., and Vitiello, M. V. (2004). Meta-analysis of quantitative sleep parameters from childhood to old age in healthy individuals: developing normative sleep values across the human lifespan. *Sleep* 27, 1255–1273. doi: 10.1093/sleep/27.7.1255

Ohrnberger, J., Fichera, E., and Sutton, M. (2017). The relationship between physical and mental health: A mediation analysis. *Soc. Sci. Med*. 195, 42–49. doi: 10.1016/j.socscimed.2017.11.008

Papp, K. V., Walsh, S. J., and Snyder, P. J. (2009). Immediate and delayed effects of cognitive interventions in healthy elderly: a review of current literature and future directions. *Alzheimers Dement*. 5, 50–60. doi: 10.1016/j.jalz.2008.10.008

Pau, M., Leban, B., Collu, G., and Migliaccio, G. M. (2014). Effect of light and vigorous physical activity on balance and gait of older adults. *Arch. Gerontol. Geriatr*. 59, 568–573. doi: 10.1016/j.archger.2014.07.008

Peralta, M., Ramos, M., Lipert, A., Martins, J., and Marques, A. (2018). Prevalence and trends of overweight and obesity in older adults from 10 European countries from 2005 to 2013. *Scand. J. Public Health*. 46, 522–529. doi: 10.1177/1403494818764810

Piercy, K. L., and Troiano, R. P. (2018). Physical activity guidelines for americans from the US department of health and human services. *Circ. Cardiovasc. Qual. Outcomes* 11:e005263. doi: 10.1161/CIRCOUTCOMES.118.005263

Posadzki, P., Pieper, D., Bajpai, R., Makaruk, H., Könsgen, N., Neuhaus, A. L., et al. (2020). Exercise/physical activity and health outcomes: an overview of Cochrane systematic reviews. *BMC Public Health* 20:1724. doi: 10.1186/s12889-020-09855-3

Ramírez-Vélez, R., Pérez-Sousa, M. A., Venegas-Sanabria, L. C., Cano-Gutierrez, C. A., Hernández-Quiñonez, P. A., Rincón-Pabón, D., et al. (2020). Normative Values for the Short Physical Performance Battery (SPPB) and their association with anthropometric variables in older colombian adults. The SABE Study, 2015. *Front. Med*. 7:52. doi: 10.3389/fmed.2020.00052

Reid, K. J., Martinovich, Z., Finkel, S., Statsinger, J., Golden, R., Harter, K., et al. (2006). Sleep: a marker of physical and mental health in the elderly. *Am. J. Geriatr. Psychiatry* 14, 860–866. doi: 10.1097/01.JGP.0000206164.56404.ba

Rosa, M. V., Perracini, M. R., and Ricci, N. A. (2019). Usefulness, assessment and normative data of the functional reach test in older adults: A systematic review and meta-analysis. *Arch. Gerontol. Geriatr.* 81, 149–170. doi: 10.1016/j.archger.2018.11.015

Rudisch, J., Jöllenbeck, T., Vogt, L., Cordes, T., Klotzbier, T. J., Vogel, O., et al. (2021). Agreement and consistency of five different clinical gait analysis systems in the assessment of spatiotemporal gait parameters. *Gait Posture* 85, 55–64. doi: 10.1016/j.gaitpost.2021.01.013

Ruegsegger, G. N., and Booth, F. W. (2018). Health benefits of exercise. *Cold Spring Harb. Perspect. Med.* 11:29694. doi: 10.1101/cshperspect.a029694

Seco, J., Abecia, L. C., Echevarría, E., Barbero, I., Torres-Unda, J., Rodriguez, V., et al. (2013). A long-term physical activity training program increases strength and flexibility, and improves balance in older adults. *Rehabil. Nurs.* 38, 37–47. doi: 10.1002/rnj.64

Serrano-Checa, R., Hita-Contreras, F., Jiménez-García, J. D., Achalandabaso-Ochoa, A., Aibar-Almazán, A., and Martínez-Amat, A. (2020). Sleep quality, anxiety, and depression are associated with fall risk factors in older women. *Int. J. Environ. Res. Public Health.* 17:4043. doi: 10.3390/ijerph17114043

Sieverdes, J. C., Wickel, E. E., Hand, G. A., Bergamin, M., Moran, R. R., and Blair, S. N. (2013). Reliability and validity of the mywellness key physical activity monitor. *Clin. Epidemiol.* 5, 13–20. doi: 10.2147/CLEP.S38370

Silva, A. A., Mello, R. G. B., Schaan, C. W., Fuchs, F. D., Redline, S., and Fuchs, S. C. (2016). Sleep duration and mortality in the elderly: a systematic review with meta-analysis. *BMJ Open* 6:e008119. doi: 10.1136/bmjopen-2015-008119

Speakman, J. R., and Westerterp, K. R. (2010). Associations between energy demands, physical activity, and body composition in adult humans between 18 and 96 y of age. *Am. J. Clin. Nutr.* 92, 826–834. doi: 10.3945/ajcn.2009.28540

Tan, X., van Egmond, L. T., Cedernaes, J., and Benedict, C. (2020). The role of exercise-induced peripheral factors in sleep regulation. *Mol. Metabo.* 42:101096. doi: 10.1016/j.molmet.2020.101096

Tang, N. K. Y., Fiecas, M., Afolalu, E. F., and Wolke, D. (2017). Changes in sleep duration, quality, and medication use are prospectively associated with health and well-being: analysis of the UK household longitudinal study. *Sleep* 40:zsw079. doi: 10.1093/sleep/zsw079

Tarazona-Santabalbina, F. J., Gómez-Cabrera, M. C., Pérez-Ros, P., Martínez-Arnau, F. M., Cabo, H., Tsaparas, K., et al. (2016). A multicomponent exercise intervention that reverses frailty and improves cognition, emotion, and social networking in the community-dwelling frail elderly: a randomized clinical trial. *J. Am. Med. Dir. Assoc.* 17, 426–433. doi: 10.1016/j.jamda.2016.01.019

Thomann, A. E., Berres, M., Goettel, N., Steiner, L. A., and Monsch, A. U. (2020). Enhanced diagnostic accuracy for neurocognitive disorders: a revised cut-off approach for the Montreal Cognitive Assessment. *Alz. Res. Therapy* 12:39. doi: 10.1186/s13195-020-00603-8

Thompson, M., and Medley, A. (2007). Forward and lateral sitting functional reach in younger, middle-aged, and older adults. *J. Geriatr. Phys. Ther.* 30:43.

Trautwein, S., Maurus, P., Barisch-Fritz, B., Hadzic, A., and Woll, A. (2019). *Recommended Motor Assessments Based on Psychometric Properties in Individuals With Dementia: A Systematic Review*. Available online at: https://link.springer.com/article/10.1186/s11556-019-0228-z (accessed November 26, 2020).

Trieu, E., and Alessi, C. (2019). "Sleep and healthy aging," in *Healthy Aging: A Complete Guide to Clinical Management*, ed P. P Coll (Cham: Springer International Publishing). doi: 10.1007/978-3-030-06200-2_23

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., and Mcdowell, M. (2008). Physical activity in the United States measured by accelerometer. *Med. Sci. Sports Exerc.* 40, 181–188. doi: 10.1249/mss.0b013e31815a51b3

Uemura, K., Shimada, H., Makizako, H., Doi, T., Yoshida, D., Tsutsumimoto, K., et al. (2013). Cognitive function affects trainability for physical performance in exercise intervention among older adults with mild cognitive impairment. *Clin. Interv. Aging.* 8, 97–102. doi: 10.2147/CIA.S39434

Vlietstra, L., Hendrickx, W., and Waters, D. L. (2018). Exercise interventions in healthy older adults with sarcopenia: A systematic review and meta-analysis. *Austr. J. Ageing* 37, 169–183. doi: 10.1111/ajag.12521

Vogel, O., Niederer, D., Wilke, J., El-Rajab, I., and Vogt, L. (2021a). Habitual physical activity and sleep duration in institutionalized older adults. *Front. Neurol.* 12:706340. doi: 10.3389/fneur.2021.706340

Vogel, O., Niederer, D., Wilke, J., Steinmann, M., and Vogt, L. (2021b). Relating lifetime activity behavior to the current level of physical activity of older adults. *J. Aging Phys. Act.* 1, 1–7. doi: 10.1123/japa.2020-0108

Vogel, O., Steinmann, M., Wilke, J., Engeroff, T., and Vogt, L. (2020). Gütekriterien der Fragebogenerfassung individueller Bewegungshistorien: Validität und Reliabilität des lifetime leisure physical activity questionnaire bei senioren in betreuten wohnanlagen. *Z. Gerontol. Geriatr.* 53, 446–450. doi: 10.1007/s00391-019-01678-1

Whibley, D., AlKandari, N., Kristensen, K., Barnish, M., Rzewuska, M., Druce, K. L., et al. (2019). Sleep and pain: a systematic review of studies of mediation. *Clin. J. Pain.* 35, 544–558. doi: 10.1097/AJP.0000000000000697

Williams, N. (2017). The borg rating of perceived exertion (RPE) scale. *Occup. Med.* 67, 404–405. doi: 10.1093/occmed/kqx063

Winter, D. A. (2009). *Biomechanics and Motor Control of Human Movement. 4th ed.* Hoboken, NJ: Wiley.

Zhang, C., Zhang, H., Zhao, M., Li, Z., Cook, C. E., Buysse, D. J., et al. (2020). Reliability, validity, and factor structure of Pittsburgh sleep quality index in community-based centenarians. *Front. Psychiatry* 11:573530. doi: 10.3389/fpsyt.2020.573530

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 30 November 2021
doi: 10.3389/fnagi.2021.722830

Effect of Fear of Falling on Mobility Measured During Lab and Daily Activity Assessments in Parkinson's Disease

Arash Atrsaei[1], Clint Hansen[2], Morad Elshehabi[2], Susanne Solbrig[3,4], Daniela Berg[2,3,4], Inga Liepelt-Scarfone[3,4,5], Walter Maetzler[2] and Kamiar Aminian[1]*

[1] Laboratory of Movement Analysis and Measurement, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, [2] Department of Neurology, UKSH, Christian-Albrechts-University, Kiel, Germany, [3] Department of Neurodegeneration, Center for Neurology and Hertie-Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany, [4] German Center for Neurodegenerative Diseases, Tübingen, Germany, [5] IB-Hochschule, Stuttgart, Germany

In chronic disorders such as Parkinson's disease (PD), fear of falling (FOF) is associated with falls and reduced quality of life. With inertial measurement units (IMUs) and dedicated algorithms, different aspects of mobility can be obtained during supervised tests in the lab and also during daily activities. To our best knowledge, the effect of FOF on mobility has not been investigated in both of these settings simultaneously. Our goal was to evaluate the effect of FOF on the mobility of 26 patients with PD during clinical assessments and 14 days of daily activity monitoring. Parameters related to gait, sit-to-stand transitions, and turns were extracted from IMU signals on the lower back. Fear of falling was assessed using the Falls Efficacy Scale-International (FES-I) and the patients were grouped as with (PD-FOF+) and without FOF (PD-FOF−). Mobility parameters between groups were compared using logistic regression as well as the effect size values obtained using the Wilcoxon rank-sum test. The peak angular velocity of the turn-to-sit transition of the timed-up-and-go (TUG) test had the highest discriminative power between PD-FOF+ and PD-FOF− (r-value of effect size = 0.61). Moreover, PD-FOF+ had a tendency toward lower gait speed at home and a lower amount of walking bouts, especially for shorter walking bouts. The combination of lab and daily activity parameters reached a higher discriminative power [area under the curve (AUC) = 0.75] than each setting alone (AUC = 0.68 in the lab, AUC = 0.54 at home). Comparing the gait speed between the two assessments, the PD-FOF+ showed higher gait speeds in the capacity area compared with their TUG test in the lab. The mobility parameters extracted from both lab and home-based assessments contribute to the detection of FOF in PD. This study adds further evidence to the usefulness of mobility assessments that include different environments and assessment strategies. Although this study was limited in the sample size, it still provides a helpful method to consider the daily activity measurement of the patients with PD into clinical evaluation. The obtained results can help the clinicians with a more accurate prevention and treatment strategy.

Keywords: inertial sensor, wearables, sit-to-stand, gait, turning, timed-up and go

OPEN ACCESS

Edited by:
Thomas Wisniewski, NYU Grossman School of Medicine, United States

Reviewed by:
Svetlana Tomic, Osijek Clinical Hospital Center, Croatia
Sachchida Nand Rai, University of Allahabad, India

***Correspondence:**
Morad Elshehabi
morad.elshehabi@uksh.de

Received: *09 June 2021*
Accepted: *21 October 2021*
Published: *30 November 2021*

Citation:
Atrsaei A, Hansen C, Elshehabi M, Solbrig S, Berg D, Liepelt-Scarfone I, Maetzler W and Aminian K (2021) Effect of Fear of Falling on Mobility Measured During Lab and Daily Activity Assessments in Parkinson's Disease. Front. Aging Neurosci. 13:722830. doi: 10.3389/fnagi.2021.722830

INTRODUCTION

Parkinson's disease (PD) is a neurodegenerative disease that is associated with the degeneration of the dopaminergic nerve cells in the substantia nigra (Rai and Singh, 2020). Although currently, there is no cure for PD, treatment options such as Levodopa focus on alleviating PD symptoms. Fear of falling (FOF) is one of the most stressful symptoms for patients with PD (Frazier, 2000; Jonasson et al., 2018), leading to reduced quality of life and social isolation (Howcroft et al., 2013). Moreover, it is the strongest predictor currently known for future falls in this population (Lindholm et al., 2015), which indirectly but strongly associates FOF with the consequence of falls, such as fractures and other injuries (Bloem et al., 2001; Allen et al., 2013).

Fear of falling can be assessed by several scales of which the Falls Efficacy Scale-International (FES-I) is the most widely used to evaluate the concerns of patients about falling during various daily activities (Delbaere et al., 2010). These activities include walking, postural transitions, and turnings during daily activities. Being subjective in nature, FOF can have impacts on mobility that can be measured objectively (Rochat et al., 2010). Therefore, by the assessment of mobility, future falls can be predicted (Delbaere et al., 2004). Inertial measurement units (IMUs) enable the objective evaluation of mobility performance, both during functional tests in the lab and during daily activities. Instrumenting functional tests such as the timed-up-and-go (TUG) and five-time sit-to-stand (5xSTS) with IMUs provide a more in-depth analysis of gait and balance performance (Salarian et al., 2010; Van Lummel et al., 2016). Furthermore, IMUs can also help clinicians to evaluate the performance of patients during daily activities that are often very different from the supervised assessment in the lab and the clinic (Warmerdam et al., 2020).

The potential of IMUs to distinguish patients with falls from those without has already been shown (Howcroft et al., 2013). These studies suggest that the most promising mobility parameters to detect an increased risk of falling are in the area of gait (Marschollek et al., 2009; Greene et al., 2010; Weiss et al., 2011, 2013), postural transition (Najafi et al., 2002; Narayanan et al., 2008; Doheny et al., 2011; Weiss et al., 2011), and turning (Haertner et al., 2018). However, none of these studies investigated the contribution of FOF to these associations in detail.

In community-dwelling older adults, it has been shown that IMU-derived TUG parameters, such as total duration, turning velocity, and sit-to-stand duration, have a significant association with the FES-I total score (Williams and Nyman, 2018). Moreover, it has been shown in patients with PD that FOF affects their turning performance during the TUG test (Haertner et al., 2018). Patients with PD with FOF had significantly lower turning peak angular velocity, and PD fallers had significantly lower gait speed, compared to non-fallers (Latt et al., 2009). A drawback of the previous studies is that the performance of the participants has been studied mostly during assessments performed in the clinic while the association between FOF and the performance of the investigated cohorts during daily activities remains unknown. This is an enormous disadvantage, as daily activity and mobility are influenced by psychological and environmental factors that cannot be effectively investigated in a supervised environment (Owsley and McGwin, 2004; Feltz and Payment, 2005; Rudman et al., 2006; Kaspar et al., 2015; Evers et al., 2020; Shah et al., 2020b; Del Din et al., 2021; Maetzler et al., 2021).

Based on these findings, the first goal of this study was to determine whether there exist mobility differences between patients with PD with (PD-FOF+) and without FOF (PD-FOF−). For this purpose, we compared the IMU-derived gait, sit-to-stand, and turning parameters from the respective lab and daily activity assessments. The second goal was to determine whether daily activity assessment can complement lab assessment in differentiating PD-FOF+ from PD-FOF−. The third goal was to investigate the associations between the same parameters obtained during these two assessment settings and study their differences in PD-FOF+ and PD-FOF−.

MATERIALS AND METHODS

Participants and Study Cohort

Twenty-six participants with PD were included in the analysis. The inclusion criteria were age between 50 and 85 years, PD based on the United Kingdom Brain-Bank Society criteria, and the ability to understand and communicate well with the investigator. Patients with dementia were excluded from the study (Emre et al., 2007). All participants gave their written informed consent and the study was approved by the ethics committee of the Medical Faculty of the University of Tübingen (protocol no. 686/2013BO1) (Haertner et al., 2018).

Lab Assessments

Lab assessments were performed during ON medication state and included the Unified Parkinson's Disease Rating Scale (UPDRS-III) (Goetz et al., 2008) and the Hoehn and Yahr (H&Y) score (Hoehn and Yahr, 2001). Fear of falling was assessed with the FES-I (Yardley et al., 2005). An FES-I score > 19 was defined as the presence of FOF (Delbaere et al., 2010). The patients were also evaluated for depressive symptoms (Beck's Depression Inventory, BDI), the amount of Levodopa equivalent dose, and quality of life (Parkinson's Disease Questionnaire, PDQ-39).

For the mobility assessments, the participants were equipped with Mobility Lab® (APDM, Portland, United States) IMUs on the lower back and the two feet. The sampling frequency was set at 128 Hz. For the analysis, accelerometer and gyroscope data were used. All participants performed first a 7-m TUG test at their convenient speed. The TUG test includes a sit-to-stand movement, a walking phase, a 180° turn, a walking back phase, and a turn-to-sit movement. The turn-to-sit transition consists of a simultaneously performed stand-to-sit transition and a 180° turn. Then, the participants performed the 5xSTS test once with their preferred speed and once as fast as possible. Rest periods were given between these three lab mobility tests.

For the analysis of the TUG test, the lower back IMU was used to analyze the sit-to-stand and stand-to-sit postural transitions with a previously validated algorithm (Atrsaei et al., 2020). The

beginning of the sit-to-stand ($t_{b,SiSt}$) and the end of the stand-to-sit ($t_{e,StSi}$) times, as well as the sit-to-stand peak power (P_{TUG}) were extracted. The two turns within the TUG were analyzed by another validated algorithm, using the data from the lower back IMU (Salarian et al., 2010). The end of the second turn ($t_{e,Turn2}$), as well as the maximum angular velocities around the vertical axis of each of the two turns ($\omega_{TUG,1}$ and $\omega_{TUG,2}$) were extracted. The total time of the TUG was calculated by subtracting the start of the sit-to-stand from the maximum value between the end of the second turn and the end of stand-to-sit: $T_{TUG} = \text{Max}\left(t_{e,StSi}, t_{e,Turn2}\right) - t_{b,SiSt}$.

The IMUs on the lower back and feet were used to extract the instantaneous gait speed during the TUG test based on the algorithm introduced in Atrsaei et al. (2021b). The mean gait speed of the whole test was calculated ($V_{TUG,avg}$).

The 5xSTS tests were analyzed by the algorithm given in Atrsaei et al. (2020), using the data obtained from the lower back IMU. The following parameters were calculated: total time and mean sit-to-stand peak power of the normal ($T_{5xSTS,N}$, $P_{5xSTS,N}$) and the fast 5xSTS ($T_{5xSTS,F}$, $P_{5xSTS,F}$).

Mobility Assessment During Daily Activities

The participants were equipped with a RehaGait® IMU (Hasomed, Magdeburg, Germany) in an elastic belt on the lower back and were asked to wear the system for 14 days. The patients were instructed to plug the sensor into a personal computer to charge during the night. The following morning, the patients were asked to unplug the sensor and wear it. The data recording was started automatically right after the sensor is unplugged. Measurement phases of less than 6 h/day were discarded from the analysis. The following mobility parameters were extracted for each patient.

Gait

Walking bouts were detected by the algorithm introduced in Atrsaei et al. (2021b). Instantaneous gait speed, i.e., gait speed at each second was calculated (Atrsaei et al., 2021b). Instances in which the gait speed was less than 0.2 m/s were not included in the walking bouts as these instances can be considered as "non-gait" periods (Atrsaei et al., 2021a). Walking bouts of less than 15 s were excluded from the analysis, to have a more steady-state gait and prevent non-locomotion movements to be detected (Atrsaei et al., 2021a). The total duration of walking for each day was obtained and was expressed as the percentage of the measurement duration of the respective day. Over all the days of measurement, the minimum ($Gait_{ALL,min}$), average ($Gait_{ALL,avg}$), and maximum ($Gait_{ALL,max}$) values of the walking percent were calculated. For instance, when a participant was assessed over a period of 5 days, and walked 5, 10, 15, 20, and 25% of the entire daily assessment periods, respectively, the $Gait_{ALL,min}$, $Gait_{ALL,avg}$, and $Gait_{ALL,max}$ would be 5, 15, and 25%, respectively.

The walking bouts were divided into short (between 15 and 30 s), medium (between 30 and 60 s), and long ones (longer than 60 s). Again, the minimum, average, and maximum values of the walking percentage per day for each type of walking bout were calculated. The indices *SWB*, *MWB*, and *LWB* were used to describe short, medium, and long walking bouts.

Over all the days of measurement stacked together, the gait speed distribution during all the walking bouts (V_{ALL}), as well as during the short (V_{SWB}), medium (V_{MWB}), and long (V_{LWB}) walking bouts were obtained separately. For each of these four distributions, the median, and the 95th percentile values were calculated.

There is evidence in the literature that gait speed often has a bimodal distribution during daily activities (Van Ancum et al., 2019; Atrsaei et al., 2021a). The first mode represents the lower preferred gait speed of the participants while the second mode represents the higher preferred gait speed of the participants (Van Ancum et al., 2019). Therefore, we also extracted the first and second modes of V_{ALL} distribution as V_{μ_1} and V_{μ_2}, respectively.

Sit-to-Stand Transitions

Sit-to-stand transitions were detected during daily activities with a validated algorithm (Atrsaei et al., 2020). For each day, the number of sit-to-stands per hour was obtained. The minimum ($SiSt_{min}$), average ($SiSt_{avg}$), and maximum ($SiSt_{max}$) number of sit-to-stands per hour were calculated over all days of measurement. Furthermore, for each sit-to-stand, the vertical peak power was determined as this parameter is a predictor of prospective falls (Regterschot et al., 2014). The distribution of all the peak power values over all the days of measurement stacked together was obtained as P_H. The median of this distribution ($P_{H,P50}$) and its 95th percentile ($P_{H,P95}$) were calculated.

Turns

Turns were detected during daily activities with a validated algorithm (El-Gohary et al., 2014). The number of turns per hour was determined for each day. The minimum ($Turns_{min}$), average ($Turns_{avg}$), and maximum ($Turns_{max}$) number of turns per hour was also calculated over all the days of measurement. For each turn, the peak angular velocity around the vertical direction was obtained. The distribution of all the peak angular velocity values over all the days of measurement stacked together were obtained as ω_H. The median ($\omega_{H,P50}$) and 95th percentile ($\omega_{H,P95}$) of this distribution were calculated.

Comparison Between PD-FOF+ and PD-FOF−

All the mobility parameters extracted from the lab and daily activity assessments were compared between PD-FOF+ and PD-FOF−. To exclude the potential differences due to gender and PD stage, the values were adjusted for gender and UPDRS-III with a multivariable logistic regression model. This analysis determines the odds of being PD-FOF+ considering gender (binary value, 0 for male, 1 for female), UPDRS-III (real-valued), and one of the mobility parameters (real-valued) explained in the previous section as independent variables. Moreover, the effect size (ES) (obtained by the *r*-value) was obtained by dividing the Wilcoxon rank-sum test statistics by the square root of the population (Ivarsson et al., 2013). An *r* value of about 0.1 indicates a small, 0.3 a medium, and 0.5 a large effect size, respectively (Cohen, 1992).

Fear of Falling Classification

To determine the predictive power of the extracted parameters in classifying PD-FOF+ and PD- FOF−, three classifiers based on a decision tree were used. Each classifier was trained based on one of the three sets of features mentioned below:

- $\mathbb{F}1$, Lab and daily activity (selected features): From all the parameters extracted from the lab and daily activity measurements, we selected those with an absolute r value of higher than 0.2. A backward elimination method was further applied to select the optimal features (Dadashi et al., 2014).
- $\mathbb{F}2$, Lab: From the set $\mathbb{F}1$, the parameters from the lab assessment were used.
- $\mathbb{F}3$, Daily activity: From the set $\mathbb{F}1$, the parameters from daily activity assessment were used.

The decision tree approach was used due to its proven performance in classifying patient populations based on mobility biomarkers (Millor et al., 2017; Rehman et al., 2019). For all the three sets mentioned above, cross-validation was performed based on the leave-one subject-out approach. The classification performance was evaluated by sensitivity, specificity, precision, accuracy, and area under the receiver operating characteristic curve (AUC) metrics.

Lab Versus Daily Activity Assessment

For each of the two groups, the gait speed, sit-to-stand peak power, and peak angular velocity were compared between lab and daily activities. For each parameter, a paired comparison was performed with the Wilcoxon sign rank test, and the significance level was set at $p = 0.05$. Pearson's correlation coefficient (ρ) was also obtained. A correlation coefficient < 0.5 was considered as low, between 0.5 and 0.7 as moderate, and >0.7 as high (Mukaka, 2012).

Moreover, each parameter obtained during the daily activities was divided by the same parameter obtained during the lab assessment. The new unitless parameters were compared between PD-FOF+ and PD-FOF− by the Wilcoxon rank-sum test.

RESULTS

Comparison Between PD-FOF+ and PD-FOF−

The characteristics of the participants are shown in **Table 1**. Of the 26 participants, nine had an FES-I score $>$ 19. The PD-FOF+ showed a trend toward higher UPDRS-III scores in comparison with the PD-FOF−. The Levodopa equivalent dose, as well as the BDI score, was not significantly different between the two groups. However, the PDQ scale, as well as its mobility subpart, were significantly different between PD-FOF+ and PD-FOF− ($p < 0.001$ and = 0.0014, respectively).

Table 2 presents results from the comparison of the lab and daily activity mobility parameters between PD-FOF+ and PD-FOF−. After the adjustment for UPDRS-III and gender, PD-FOF+ participants had significantly longer T_{TUG} accompanied by slower $\omega_{TUG,1}$, $\omega_{TUG,2}$, and a slower $V_{TUG,avg}$ which, however, did not reach significance.

TABLE 1 | Comparison of PD-FOF+ and PD-FOF−.

Parameter	PD-FOF+	PD-FOF−	*p*-value	ES
Number	9 (9 females)	17 (12 females)	0.083	0.35
Age (year)	65 [62, 69]	64 [58, 75]	0.829	0.05
Height (m)	1.78 [1.67, 1.83]	1.75 [1.69, 1.79]	0.608	0.11
Weight (kg)	81.0 [77.0, 86.0]	77 [70.5, 97.0]	0.935	−0.02
UPDRS-III (0-132)	30 [24, 34]	22 [18, 28]	0.053	0.38

The p-value was obtained by Wilcoxon rank sum test. Significance level was set at 0.05. Except the number of participants, the values are shown with Median [IQR]. ES is the effect size obtained by the r-value.

Several parameters were slightly different between the two populations although the logistic regressions showed no statistical significance. For instance, compared with PD- FOF−, PD-FOF+ had, on average, lower gait speeds during the TUG ($V_{TUG,avg}$) and daily activities (V_{μ_1}), longer $T_{5xSTS,F}$, lower percentages of walking bouts (i.e., $Gait_{ALL,min}$, $Gait_{ALL,avg}$, and $Gait_{ALL,max}$), and lower numbers of sit-to-stands ($SiSt_{max}$) and turns ($Turns_{min}$) per hour during daily activities.

No significant differences were found between the two groups when dividing the walking bouts based on their duration (**Table 3**). However, PD-FOF+ tended to have a lower percentage of short (e.g., $Gait_{SWB,max}$) and long (e.g., $Gait_{LWB,max}$) walking bouts, compared with PD-FOF− (**Table 3**).

The effect sizes of the parameters are shown in **Table 2** and **Figure 1** in descending order. As a general note, lab-extracted parameters showed higher effect sizes than those extracted from the daily activity assessment. $\omega_{TUG,2}$ had the highest effect size, followed by other parameters extracted from the TUG test (except P_{TUG} which had a very small effect size, see also **Figure 1**). $\omega_{TUG,1}$ had a lower effect size than $\omega_{TUG,2}$. Directly after the TUG test, parameters ranked the $T_{5xSTS,F}$ and $P_{5xSTS,F}$ from the 5xSTS test with fast speed. The effect sizes of the parameters from the 5xSTS with normal speed ($T_{5xSTS,N}$ and $P_{5xSTS,N}$) were lower than those from the fast version. $T_{5xSTS,N}$ had a smaller effect size compared with $P_{5xSTS,N}$. $Gait_{ALL,max}$, $SiSt_{max}$, $Turns_{min}$, and V_{μ_1} had the highest effect sizes among the daily activity parameters, and the median gait speed ($V_{ALL,P50}$) the lowest.

Fear of Falling Classification

Out of the 41 mobility parameters, 23 had an effect size $>$ 0.2 (**Figure 1**). From these parameters, 19 features (used for machine learning-based classifier; marked with ^{x} in **Figure 1**) were selected by the backward elimination method and used for the $\mathbb{F}1$ set.

Based on the three sets of features mentioned in the "Fear of Falling Classification" section, the results of the classification are shown in **Table 4**. The best performance was achieved based on set $\mathbb{F}1$ which was a combination of features obtained from the lab and daily activity assessments (AUC = 0.75). The accuracy of this set was higher than when using lab ($\mathbb{F}2$, AUC = 0.68) or daily activity ($\mathbb{F}3$, AUC = 0.54) features alone.

TABLE 2 | Comparison of the extracted parameter between the group with fear of falling (FOF) (PD-FOF+) and without (PD-FOF−).

Category	Parameter	PD-FOF+	PD-FOF−	*p*-value	ES
TUG	T_{TUG} (s)	19.93 [19.29, 21.43]	17.40 [15.23, 19.38]	0.044*	0.51
	$V_{TUG,avg}$ (m/s)	1.01 [0.95, 1.08]	1.13 [1.05, 1.29]	0.069	−0.47
	$\omega_{TUG,1}$ (deg/s)	124.4 [119.1, 165.0]	161.6 [149.4, 202.0]	0.029*	−0.36
	$\omega_{TUG,2}$ (deg/s)	110.2 [103.0, 132.4]	158.5 [140.2, 167.3]	0.018*	−0.61
	P_{TUG} (W)	44.11 [16.07, 49.26]	37.32 [28.61, 45.08]	0.269	−0.01
Normal 5xSTS	$T_{5xSTS,N}$ (s)	17.02 [15.79, 21.61]	16.94 [15.05, 21.11]	0.772	−0.10
	$P_{5xSTS,N}$ (W)	44.86 [32.96, 51.29]	39.13 [32.47, 50.02]	0.949	0.33
Fast 5xSTS	$T_{5xSTS,F}$ (s)	14.24 [13.58, 15.02]	11.48 [11.03, 16.32]	0.594	0.35
	$P_{5xSTS,F}$ (W)	65.32 [46.86, 74.92]	49.16 [37.71, 70.15]	0.373	0.35
Gait at Home	$V_{ALL,P50}$ (m/s)	0.81 [0.79, 0.93]	0.88 [0.76, 0.93]	0.660	0.01
	$V_{ALL,P95}$ (m/s)	1.17 [1.11, 1.33]	1.23 [1.09, 1.27]	0.822	0.13
	V_{μ_1} (m/s)	0.49 [0.36, 0.59]	0.63 [0.43, 0.75]	0.053	−0.31
	V_{μ_2}(m/s)	0.91 [0.82, 1.00]	0.96 [0.84, 1.06]	0.447	−0.16
	$Gait_{ALL,min}$ (%)	0.47 [0.43, 0.83]	0.93 [0.62, 2.86]	0.122	−0.29
	$Gait_{ALL,avg}$ (%)	3.10 [2.69, 3.33]	4.28 [2.78, 5.95]	0.174	−0.24
	$Gait_{ALL,max}$ (%)	6.59 [4.56, 7.88]	8.05 [6.71, 13.34]	0.177	−0.33
Sit-to-stand at Home	$P_{H,P50}$ (W)	18.72 [12.62, 24.33]	19.54 [13.00, 25.72]	0.291	−0.05
	$P_{H,P95}$ (W)	43.10 [36.69, 54.54]	43.22 [33.30, 62.99]	0.239	−0.03
	$SiSt_{min}$ (/h)	1.75 [0.96, 2.60]	1.75 [0.80, 3.27]	0.998	−0.05
	$SiSt_{avg}$ (/h)	3.74 [2.80, 4.60]	4.45 [3.42, 5.17]	0.670	−0.24
	$SiSt_{max}$ (/h)	5.42 [4.11, 6.39]	6.27 [5.59, 8.40]	0.254	−0.33
Turn at Home	$\omega_{H,P50}$ (deg/s)	60.24 [58.67, 63.78]	63.55 [59.13, 68.70]	0.638	−0.23
	$\omega_{H,P95}$ (deg/s)	110.6 [107.2, 123.1]	111.1 [108.1, 123.4]	0.758	−0.03
	$Turns_{min}$ (/h)	55.71 [48.56, 68.32]	74.30 [55.86, 83.79]	0.923	−0.33
	$Turns_{avg}$ (/h)	86.46 [82.02, 96.16]	102.5 [85.81, 118.2]	0.578	−0.24
	$Turns_{max}$ (/h)	123.7 [109.4, 192.1]	142.0 [124.4, 163.0]	0.553	0.10

*The p-value shows the significance of the coefficient of the inertial measurement unit (IMU)-based parameter in the logistic regression. *P < 0.05 was considered significant. The values of IMU-based parameters are shown by Median [IQR]. The effect size obtained by the r-value is ES.*

TABLE 3 | Comparison of the extracted parameter for short, medium, and long walking bouts (WB) between PD-FOF+ and PD-FOF−.

WB	Parameter	PD-FOF+	PD-FOF−	*p*-value	ES
Short	$V_{SWB,P50}$ (m/s)	0.70 [0.63, 0.71]	0.71 [0.60, 0.76]	0.533	0.04
	$V_{SWB,P95}$ (m/s)	1.10 [1.06, 1.19]	1.07 [1.01, 1.20]	0.529	0.10
	$Gait_{SWB,min}$ (%)	0.43 [0.39, 0.49]	0.69 [0.36, 1.31]	0.187	−0.24
	$Gait_{SWB,avg}$ (%)	1.40 [1.19, 1.68]	1.72 [1.30, 2.47]	0.167	−0.26
	$Gait_{SWB,max}$ (%)	2.65 [2.19, 2.81]	3.15 [2.31, 4.50]	0.095	−0.23
Medium	$V_{MWB,P50}$ (m/s)	0.83 [0.79, 0.89]	0.85 [0.78, 0.92]	0.761	0.09
	$V_{MWB,P95}$ (m/s)	1.12 [1.08, 1.27]	1.16 [1.06, 1.24]	0.613	0.12
	$Gait_{MWB,min}$ (%)	0.00 [0.00, 0.07]	0.00 [0.00, 0.13]	0.271	−0.19
	$Gait_{MWB,avg}$ (%)	0.51 [0.29, 0.67]	0.62 [0.52, 0.92]	0.110	−0.25
	$Gait_{MWB,max}$ (%)	1.38 [0.84, 1.88]	1.57 [1.24, 2.35]	0.184	−0.22
Long	$V_{LWB,P50}$ (m/s)	0.92 [0.89, 1.05]	0.96 [0.89, 1.07]	0.859	0.03
	$V_{LWB,P95}$ (m/s)	1.24 [1.12, 1.40]	1.24 [1.11, 1.32]	0.772	0.15
	$Gait_{LWB,min}$ (%)	0.00 [0.00, 0.00]	0.00 [0.00, 0.05]	0.279	−0.16
	$Gait_{LWB,avg}$ (%)	1.08 [0.74, 1.24]	1.87 [0.78, 2.70]	0.431	−0.28
	$Gait_{LWB,max}$ (%)	3.24 [2.26, 5.31]	5.43 [3.94, 7.08]	0.314	−0.31

*The p-value shows the significance of the coefficient of the IMU-based parameter in the logistic regression. *P < 0.05 was considered significant. The values of IMU-based parameters are shown by Median [IQR]. The effect size obtained by the r-value is ES.*

The sensitivity of the classification based on the features from the lab ($\mathbb{F}2$) was higher than that obtained from the daily activity features, while the specificity of the classification based on daily activity features ($\mathbb{F}3$) was higher. Moreover, $\mathbb{F}2$ features achieved higher accuracy and AUC values, than the $\mathbb{F}3$ features.

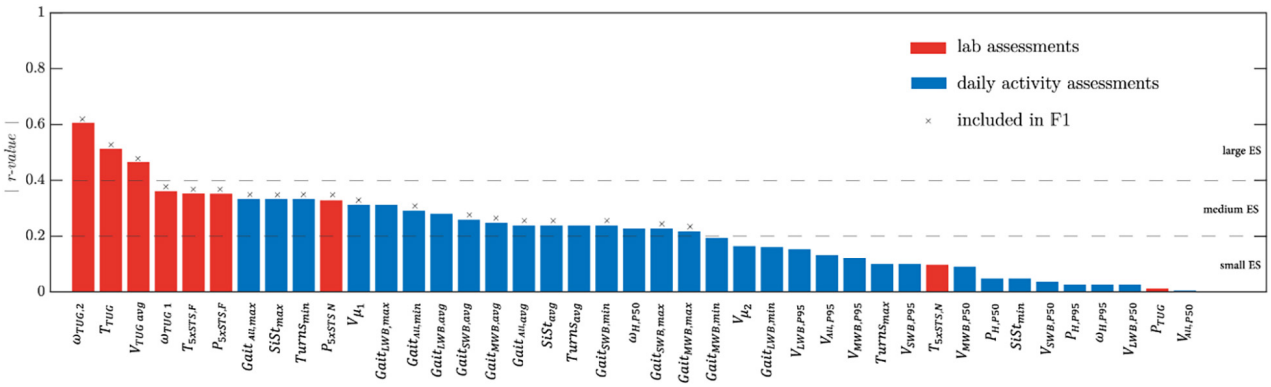


FIGURE 1 | Absolute effect size values (*r*-value) of the mobility parameters extracted from the lab (in red) and daily activity assessments (in blue). The features selected for the $\mathbb{F}1$ feature set are marked by a cross (see section "Fear of Falling Classification").

TABLE 4 | The performance metrics of the classification of PD-FOF+ vs. PD-FOF−.

Feature set	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)	AUC
$\mathbb{F}1$, Lab and daily activity	55.6	94.1	83.3	80.8	0.75
$\mathbb{F}2$, Lab	57.7	64.7	40.0	65.4	0.68
$\mathbb{F}3$, Daily activity	44.4	76.5	50.0	57.7	0.54

*$\mathbb{F}1$: Selected 19 features marked with crosses in **Figure 1**. $\mathbb{F}2$: 7 lab features from $\mathbb{F}1$. $\mathbb{F}3$: 12 daily activity features from $\mathbb{F}1$.*

Lab Versus Daily Activity Assessment

The results of the paired comparison between lab and daily activity assessments for gait speed, sit-to-stand peak power, and turning peak angular velocity are shown in **Table 5**. In the PD-FOF+ group, no significant correlations were found between lab and daily activity assessments concerning gait speed. Moreover, PD-FOF+ had significantly higher gait speeds at the 95th percentile of their walking speed distributions compared with the lab ($V_{ALL,P95}$, $V_{SWB,P95}$, $V_{MWB,P95}$, and $V_{LWB,P95}$). In the PD-FOF− group, $V_{TUG,avg}$ had a significant but low correlation with $V_{H,P95}$ ($\rho = 0.48$). A high correlation was also observed between $V_{TUG,avg}$ and V_{μ_2} ($\rho = 0.70$). Moderate correlations were observed between $V_{TUG,avg}$ and a gait speed of medium ($V_{MWB,P95}$) and long ($V_{LWB,P50}$) walking bouts ($\rho = 0.59$ and $\rho = 0.57$, respectively). The PD-FOF− group walked significantly faster during the TUG than during their daily activities.

Regarding the sit-to-stand peak power, a high and significant correlation was found between P_{TUG} and $P_{H,P95}$ for both groups (PD-FOF+, $\rho = 0.77$; PD- FOF−, $\rho = 0.79$). In both groups, $P_{5xSTS,N}$ had a high and significant correlation with $P_{H,P95}$ (PD-FOF+, $\rho = 0.83$; PD- FOF−, $\rho = 0.70$). No significant correlations were found between the 5xSTS with fast speed and daily activity assessment. Both groups had significantly higher peak power during the 5xSTS tests compared with $P_{H,P50}$ during daily activities. However, $P_{H,P95}$ values were not significantly different from the 5xSTS tests in the lab.

Finally, for turning peak angular velocity, no significant correlations were found between the lab and daily activities in any group. For PD-FOF+, there were no significant differences between $\omega_{H,P95}$ and both turns of the TUG. However, PD-FOF− had faster turns in the lab, compared with the home environment.

For a better representation of lab versus daily activity parameters, the gait speed, sit-to-stand peak power, and turning peak angular velocity are presented in **Figure 2** as unitless ratios (daily activity parameter divided by the respective lab parameter). Most of the ratios were less than 1 (i.e., lower value of a parameter in the daily life environment). However, a few parameters, e.g., $\frac{V_{ALL,P95}}{V_{TUG,avg}}$, $\frac{P_{ALL,P95}}{P_{TUG}}$, had values > 1, preferentially in the PD-FOF+ group. Moreover, when comparing PD-FOF+ with PD-FOF−, significant differences were found for the ratios $\frac{V_{ALL,P95}}{V_{TUG,avg}}$, $\frac{V_{SWB,P95}}{V_{TUG,avg}}$, $\frac{V_{MWB,P95}}{V_{TUG,avg}}$, $\frac{V_{LWB,P50}}{V_{TUG,avg}}$, $\frac{V_{LWB,P95}}{V_{TUG,avg}}$, $\frac{\omega_{H,P50}}{V_{TUG,avg}}$, and $\frac{\omega_{H,P95}}{\omega_{TUG,2}}$, with higher ratios in the PD-FOF+ group.

DISCUSSION

Most of the previous studies on this topic that have shown mobility-associated differences between PD-FOF+ and PD-FOF− have investigated their participants only in the lab. In this study, thanks to IMUs and dedicated algorithms, several mobility parameters were collected from patients with PD with and without FOF, when performing functional tests in the lab and living in their usual environment. The effect of FOF was investigated by quantifying the changes in mobility parameters between lab and daily life. The discriminative power between PD-FOF+ and PD-FOF− was shown by a logistic regression model considering each setting separately and in combination. And finally, the association between the lab and daily activity setting was studied by considering their correlation.

TABLE 5 | Paired comparison of the parameters between lab and home.

Category	Lab	Daily activity	Difference, *p*-value		Correlation (ρ)	
			PD-FOF+	PD-FOF−	PD-FOF+	PD-FOF−
Gait speed	$V_{TUG,avg}$	$V_{ALL,P50}$	0.023*	< 0.001*	−0.34	0.36
		$V_{ALL,P95}$	0.008*	0.836	−0.15	0.48*
		V_{μ_1}	0.008*	< 0.001*	0.35	0.44
		V_{μ_2}	0.039*	< 0.001*	−0.03	0.70*
		$V_{SWB,P50}$	0.008*	< 0.001*	−0.14	0.38
		$V_{SWB,P95}$	0.039*	0.044*	0.20	0.40
		$V_{MWB,P50}$	0.016*	< 0.001*	−0.48	0.44
		$V_{MWB,P95}$	0.008*	0.309	−0.04	0.59*
		$V_{LWB,P50}$	0.148	< 0.001*	−0.12	0.57*
		$V_{LWB,P95}$	0.008*	0.193	−0.10	0.42
Sit-to-stand peak power	P_{TUG}	$P_{H,P50}$	0.039*	0.001*	0.75*	0.48
		$P_{H,P95}$	0.541	0.006*	0.77*	0.79*
	$P_{5xSTS,N}$	$P_{H,P50}$	0.015*	0.003*	0.83*	0.13
		$P_{H,P95}$	0.578	0.167	0.83*	0.70*
	$P_{5xSTS,F}$	$P_{H,P50}$	0.031*	0.002*	−0.28	0.12
		$P_{H,P95}$	0.312	0.492	−0.34	0.61
Turning peak angular velocity	$\omega_{TUG,1}$	$\omega_{H,P50}$	< 0.001*	< 0.001*	0.28	0.20
		$\omega_{H,P95}$	0.139	< 0.001*	−0.44	0.00
	$\omega_{TUG,2}$	$\omega_{H,P50}$	< 0.001*	< 0.001*	0.69	0.43
		$\omega_{H,P95}$	0.815	< 0.001*	0.57	0.00

*p-value from the Wilcoxon sign rank test and Pearson's correlation coefficient (ρ) describe the differences of the parameters between lab and daily life. The significance level was set to 0.05 and shown with *. Significant correlation coefficients were marked with *.*

Regarding the effect of FOF on mobility, PD-FOF+ needed more time to perform the TUG test than the PD- FOF−, which was –at least partly- explained by the slower performance of the two turns included in this test (**Table 2**). This supports previous findings (Bryant et al., 2014; Haertner et al., 2018; Abou et al., 2021) and suggests that PD-FOF+ suffer from increased fear, especially during turns. This fear may be justified, e.g., through increased dysbalance or other constraints associated with FOF (Pourghayoomi et al., 2020). The larger difference between the two groups in the second turn, which also includes a stand- or walk-to-sit movement, may also argue for the different balance capacities between the groups. This argument is further supported by the slower peak angular velocity during the second turn compared with the first turn in the PD-FOF+ group.

In contrast to the evidence in the literature (Bryant et al., 2014), we did not observe a significant difference in gait speed between PD-FOF+ and PD-FOF− during the TUG test ($V_{TUG,avg}$). As the *r*-value showed a large effect size for this parameter in both groups, we hypothesize that PD severity rather than FOF has a particular influence on this parameter. We performed a Wilcoxon rank-sum test on $V_{TUG,avg}$ without adjusting for the aforementioned confounders, and obtained a significant difference between the PD-FOF+ and PD-FOF− ($p = 0.021$). Therefore, more evidence with a larger dataset is required to confirm this hypothesis as most of the previous studies did not adjust the statistical analysis for potential confounders.

Although none of the 5xSTS tests could sufficiently discriminate between PD-FOF+ and PD- FOF−, the fast 5xSTS test presented larger effect sizes than the preferred speed 5xSTS test (**Table 2** and **Figure 1**). This is an argument for including the fast version rather than the preferred speed version (Goldberg et al., 2012; Staartjes and Schröder, 2018) in the assessment panel of clinical protocols. For the 5xSTS with preferred speed, the mean peak power of sit-to-stands ($P_{5xSTS,N}$) had a medium effect size while the effect size for the total duration of the test ($T_{5xSTS,N}$) was low (**Table 2** and **Figure 1**). This again highlights the usefulness of an instrumented 5xSTS test with IMUs to extract biomechanical parameters beyond the conventionally measured duration of the test (Van Lummel et al., 2016). Nevertheless, the IMU-derived sit-to-stand peak power did not differentiate PD-FOF+ from PD-FOF−. Also, the sit-to-stand peak power derived from the TUG test (P_{TUG}) was not significantly different between the groups. An explanation can be that the PD-FOF+ group might not have particular difficulties in performing postural transitions. However, numerous studies showed the predictive power of the 5xSTS test for future falls (Buatois et al., 2008; Duncan et al., 2011; Doheny et al., 2013; Qiu et al., 2018). Therefore, our results, together with previous results, suggest that the 5xSTS test is associated with aspects of falls that are independent of FOF.

None of the parameters derived from the daily activity assessment could significantly differentiate PD-FOF+ from PD-FOF−. However, medium effect size values were observed for several parameters. Interestingly, the effect size for the lower preferred gait speed (V_{μ_1}) was higher than the median or 95th percentile values of gait speed distribution. This

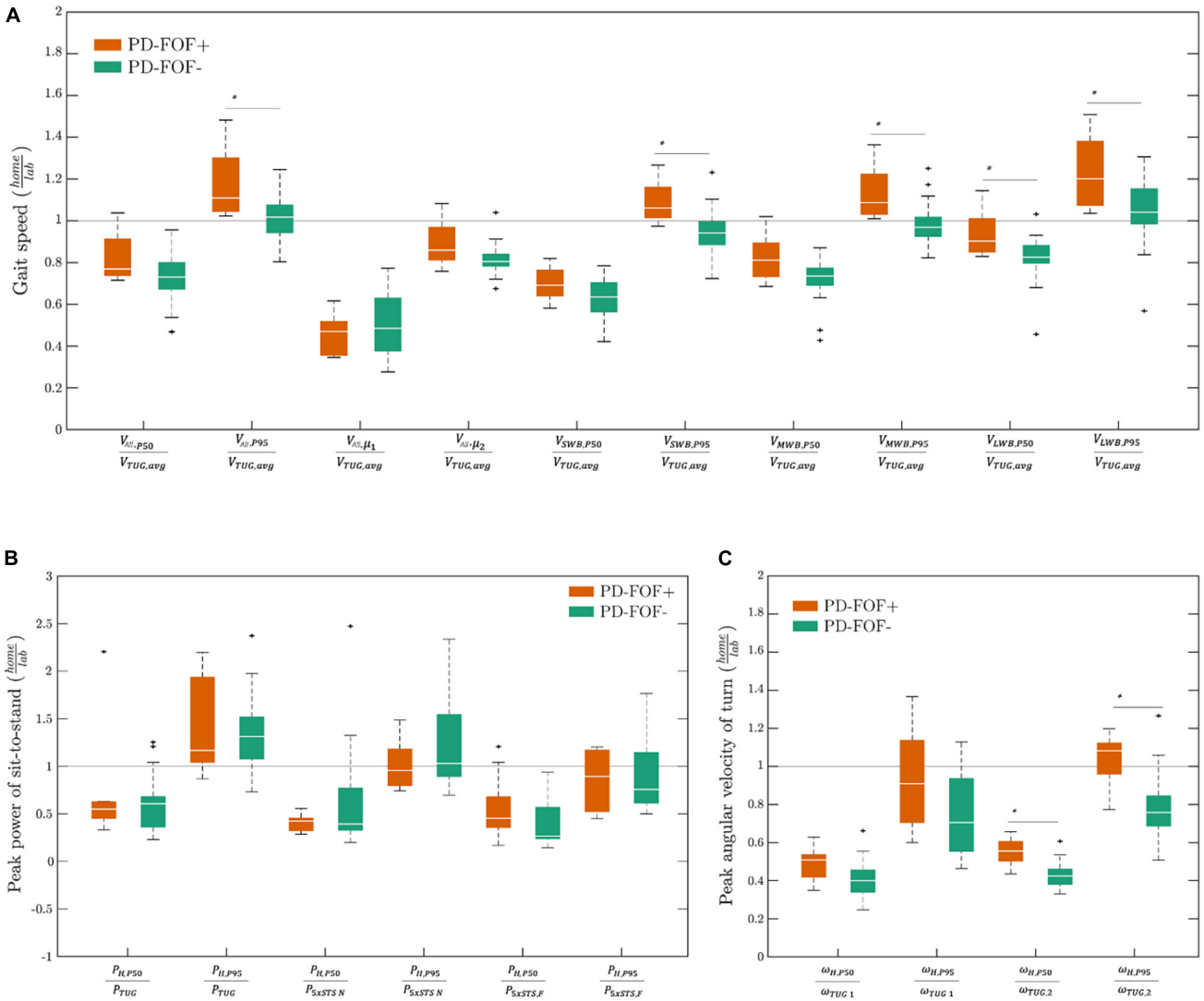


FIGURE 2 | Unitless daily activity divided by lab parameter ratios of **(A)** gait speed, **(B)** sit-to-stand peak power, and **(C)** turning peak angular velocity in PD-FOF+ and PD-FOF−. Differences between the groups were analyzed by the Wilcoxon rank-sum test. Only significant differences were shown on the plots $^*p < 0.05$.

shows the importance of more precise modeling of gait speed distribution, rather than assuming a simple normal distribution of the obviously complex movements that occur in the usual environment (which was done in most of the previous studies, e.g., Toosizadeh et al., 2015; Takayanagi et al., 2019; Shah et al., 2020b). Interestingly, V_{μ_1} showed a higher effect size than V_{μ_2}. It should be noted that V_{μ_1} is assumed to correspond more to shorter walking bouts and V_{μ_2} represents mostly longer walking bouts that are more likely to occur outdoors (Van Ancum et al., 2019). Thus, our results regarding the higher effect size of V_{μ_1} vs. V_{μ_2} suggest that shorter walking bouts are more meaningful to describe mobility performance (limitations) of PD-FOF+, and maybe an interesting therapeutic target for future trials. It could also be speculated that PD-FOF+ have more problems than PD-FOF− during multitask-walking, as shorter walking bouts have obviously a higher probability to be associated with additional tasks, compared with long walking bouts which have a high probability for reflection, e.g., walks without relevant dual-task claim. Therefore, according to **Figure 1**, it is not surprising that the features that remained for the classification included more parameters from short walking bouts ($Gait_{SWB,min}$, $Gait_{SWB,avg}$, and $Gait_{SWB,max}$) than from medium and long walking bouts ($Gait_{MWB,avg}$ and $Gait_{MWB,max}$).

In addition to $Gait_{ALL,max}$, $SiSt_{max}$ and $Turn_{min}$ were among the daily activity parameters with the highest effect sizes. Thus, the number of various types of activities should also be considered in addition to parameters such as gait speed, sit-to-stand peak power, and turning peak angular velocity that characterizes these activities. Moreover, there was a tendency toward a lower amount of activity in PD-FOF+. The PDQ

score was significantly lower in PD-FOF+, showing that the quality of life of the patients was highly affected by their FOF. This can explain the lower amount of daily activities in this group of patients.

After feature selection in section "Lab Versus Daily Activity Assessment," several parameters from the lab and daily activity assessments remained in the selected features (**Figure 1**). Training three classifiers based on three sets of features, i.e., $\mathbb{F}1$, $\mathbb{F}2$, and $\mathbb{F}3$, revealed that set $\mathbb{F}1$ led to the most accurate classifier to distinguish the PD-FOF+ from the PD-FOF− group (**Table 4**). This selection, including features from both the lab and daily activity assessments, further supports the usefulness of including daily activity assessments in clinical practice as they have complementary information to the assessments performed in the lab (Maetzler et al., 2021). The more accurate classification of FOF with lab features ($\mathbb{F}2$), compared with daily activity features ($\mathbb{F}3$, **Table 4**), suggests that capacity aspects play an important role for the definition of FOF (Maetzler et al., 2021) and functional tests in the lab should always be performed for the evaluation in FOF. Still, the inclusion of environmental context and psychological factors from daily life is a valuable addition and can contribute to increased specificity.

Comparing the gait speed between the lab and daily activity assessments, significant correlations were found for PD-FOF− but not PD-FOF+ (**Table 5**). Interestingly, PD-FOF+ had higher gait speed values in the "capacity" area of their daily activity assessment compared with the lab. For these participants, $\frac{V_{P95}}{V_{TUG,avg}}$, $\frac{V_{SWB,P95}}{V_{TUG,avg}}$, $\frac{V_{MWB,P95}}{V_{TUG,avg}}$, and $\frac{V_{LWB,P95}}{V_{TUG,avg}}$ had values greater than 1 (**Table 5**). One explanation can be that PD-FOF+ might be more cautious in non-familiar environments such as the lab. Moreover, and potentially more relevant for future management strategies, they might have been less cautious in their daily life especially when it comes to fast (and therefore more dangerous) gait episodes (Salkovic et al., 2017).

Another interesting observation, in our view, was that in PD-FOF−, $V_{TUG,avg}$ was significantly correlated with parameters during daily activity assessments that represent mostly the capacity aspects, i.e., $V_{H,P95}$, V_{H,μ_2}, $V_{MWB,P95}$, and $V_{LWB,P50}$. Moreover, the correlation between $V_{TUG,avg}$ and $V_{\mu_2}$was high ($\rho = 0.70$). These findings firstly confirm the relevant association of lab parameters with daily activity parameters that are near the capacity area (Van Ancum et al., 2019; Warmerdam et al., 2020). These results suggest that capacity-associated values obtained during daily activities can indeed predict the capacity of a participant in the lab. Furthermore, the high association between $V_{TUG,avg}$ and V_{μ_2} is again in favor of considering a bimodal gait speed distribution during daily activities (Atrsaei et al., 2021a).

Regarding the sit-to-stand peak power, $P_{H,P95}$ had high correlations with P_{TUG} and $P_{5xSTS,N}$ but not with $P_{5xSTS,F}$ (**Table 5**). This indicates that the 5xSTS test with preferred speed and the sit-to-stand part of the TUG test is most representative of the sit-to-stands performed during daily activities. In fact, in the TUG test, it is more accurate to name the initial postural transition as sit-to-walk rather than sit-to-stand. Since in daily life, there is often more sitting-to-walking than sitting-to-standing, the high correlation between P_{TUG} and $P_{H,P95}$ seems reasonable. Therefore, to have a better understanding of the sit-to-stand performance of patients during daily activities, clinicians should consider the 5xSTS test with preferred speed and the TUG sit-to-stand movement, rather than the fast 5xSTS test. The high association of sit-to-stand peak power between the lab and daily activity assessments was also observed in a study in community-dwelling older adults (Zhang et al., 2017). Nevertheless, as we demonstrated earlier, the 5xSTS test with fast speed had higher discriminative power for differentiating PD-FOF+ from PD-FOF−.

Our results are comparable to a very recent study on the impact of FOF on mobility parameters in a relatively large population of community-dwelling older adults (Wang et al., 2021). In that study, FOF led to a poorer mobility performance during both lab and daily activity assessments. Moreover, and comparable to this study, the consideration of both assessments showed the best discriminatory power between the presence and absence of FOF (lab assessment, AUC = 0.64; lab and daily activity assessment, AUC = 0.77). The strengths of our study, compared with the aforementioned study, are that we included postural transition and turning in addition to walking, and we assessed the daily activity over an average period of 12 (and not only 2) days (Wang et al., 2021).

Our study faces some limitations. First, our sample size could be small for statistical analyses. The observations and results could be supported more strongly with a larger population. This could explain why the parameters obtained during the daily activities did not differ significantly between PD-FOF+ and PD-FOF−. For instance, V_{μ_1} was at the edge of a statistically significant difference. However, it should be noted that finding participants with a specific impairment that are willing to participate in several clinical assessments, as well as 2 weeks of activity monitoring, can be challenging. While in this study, we explored the difference between participants with low and moderate FOF, the difference between participants with low and high FOF might be more evident with mobility parameters obtained during daily activities. Using other questionnaires in addition to FES-I can also be investigated. For example, participants can be asked whether their FOF restricts their activities or not (Rochat et al., 2010).

The duration of daily activities measurements could still be increased to have a more accurate estimation of the daily routines of the patients. Nevertheless, considering the current usability of IMUs, especially for older adults, it is a bit challenging to engage the participants for more than 2 weeks of measurements. It is not surprising that in the literature, a lot of studies consider only 1 week of daily activity measurements (Storm et al., 2018; Galperin et al., 2019; Van Ancum et al., 2019; Shah et al., 2020a).

Another point of limitation can be the turning assessment during daily activities. The turning algorithm considered turns with durations of 0.5–10 s and angles $> 45°$ (El-Gohary et al., 2014). This is a broad range, and future studies should investigate whether more specific definitions for turns that are performed in daily life have higher discriminatory power. Furthermore, the employed algorithm detected turns regardless of their occurrence during walking or sedentary behavior. Although it might be rare, participants might have been in a sitting position in a

moving vehicle that had similar turning to those of a human that walks and turns at the same time. Therefore, further work is required to adapt the algorithm to detect turnings that occur during locomotion.

Another point was that since some of the unitless parameters showed significant differences between the two groups (**Figure 2**), we were curious if adding them to the feature set $\mathbb{F}1$ will improve the classification results in **Table 4**. However, no improvement was observed. The reason might be that the ratio of home-derived mobility parameter divided by the same parameter obtained in the lab did not bring additional information as the information regarding both of the assessment settings was already there. Finally, to keep the data accuracy as high as possible, we excluded walking bouts < 15 s from the analysis to prevent other activities from being wrongly detected as a walking bout. However, these walking bouts contribute to a relevant portion of daily walking (Del Din et al., 2016; Shah et al., 2020c), and removing them might affect the meaningfulness of walking parameters with respect to the actual research question.

To conclude, the use of the IMU along with the dedicated algorithms allowed an unobtrusive assessment of mobility during daily activities. Although lab-based mobility parameters had generally higher discriminative power in differentiating PD-FOF+ and PD- FOF−, integrating daily activity assessments provided a more accurate classification of these patients. By comparing the same parameters from both settings, we could show for the first time that (i) considering lab and daily activity mobility parameters can lead to more accurate classification of PD-FOF+ and PD-FOF− compared with each lab and daily activity assessments alone (ii) the PD-FOF+ group performs the lab assessments with a rather cautious gait but used a rather incautious gait pattern in the usual environment; (iii) the sit-to-stand peak power of the 5xSTS test with preferred speed and of the TUG was more closely associated with sit-to-stand movement in daily life, than was the same parameter obtained from the fast 5xSTS, and (iv) the 5xSTS test with fast speed mostly measured the capacity aspects of daily activities. These results provide further insight into the daily life behavior of PD patients with FOF, can stimulate prevention and treatment strategies, and can serve as a template for further studies using these novel techniques and assessment strategies.

DATA AVAILABILITY STATEMENT

The datasets generated and/or analyzed during this study are not publicly available but might be accessible from the corresponding author on reasonable request, upon joint approval from AA, CH, WM, and KA.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Medical Faculty of the University of Tübingen (protocol no. 686/2013BO1). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AA designed the technical research and scientific question, extracted and analyzed the mobility parameters from clinical and home-based measurements, and wrote the manuscript. CH, ME, SS, DB, and IL-S performed the clinical measurements and extracted and analyzed the clinical questionnaires. AA, WM, DB, and KA discussed the results. WM, DB, and KA supervised and led the study. All the authors read, revised, and approved the final manuscript.

FUNDING

This study was supported by an unrestricted grant from Janssen Research and Development, a division of Janssen Pharmaceutica N.V. The funding of the study was pre-competitive and the funders did not have any influence on the design, conduction, and analysis of the results presented in this study. We acknowledge financial support by Land Schleswig-Holstein within the funding programme Open Access Publikationsfonds.

ACKNOWLEDGMENTS

We thank all participants that took part in the study.

REFERENCES

Abou, L., Alluri, A., Fliflet, A., Du, Y., and Rice, L. A. (2021). Effectiveness of physical therapy interventions in reducing fear of falling among individuals with neurologic diseases: a systematic review and meta-analysis. *Arch. Phys. Med. Rehabil.* 102, 132–154. doi: 10.1016/j.apmr.2020.06.025

Allen, N. E., Schwarzel, A. K., and Canning, C. G. (2013). Recurrent falls in parkinson's disease: a systematic review. *Parkinsons Dis.* 2013:906274. doi: 10.1155/2013/906274

Atrsaei, A., Dadashi, F., Hansen, C., Warmerdam, E., Mariani, B., Maetzler, W., et al. (2020). Postural transitions detection and characterization in healthy and patient populations using a single waist sensor. *J. Neuroeng. Rehabil.* 17:70. doi: 10.1186/s12984-020-00692-4

Atrsaei, A., Dadashi, F., Mariani, B., Gonzenbach, R., and Aminian, K. (2021b). Toward a remote assessment of walking bout and speed: application in patients with multiple sclerosis. *IEEE J. Biomed. Heal. Inform.* 25, 4217–4228. doi: 10.1109/JBHI.2021.3076707

Atrsaei, A., Corrà, M., Dadashi, F., Vila-Chã, N., Maia, L., Mariani, B., et al. (2021a). Gait speed in clinical and daily living assessments in Parkinson's disease patients: performance versus capacity. *NPJ Parkinsons Dis.* 7:24. doi: 10.1038/s41531-021-00171-0

Bloem, B. R., Grimbergen, Y. A. M., Cramer, M., Willemsen, M., and Zwinderman, A. H. (2001). Prospective assessment of falls in Parkinson's disease. *J. Neurol.* 248, 950–958. doi: 10.1007/s004150170047

Bryant, M. S., Rintala, D. H., Hou, J. G., and Protas, E. J. (2014). Influence of fear of falling on gait and balance in Parkinson's disease. *Disabil. Rehabil.* 36, 744–748. doi: 10.3109/09638288.2013.814722

Buatois, S., Miljkovic, D., Manckoundia, P., Gueguen, R., Miget, P., Vançon, G., et al. (2008). Five times sit to stand test is a predictor of recurrent falls in healthy community-living subjects aged 65 and older. *J. Am. Geriatr. Soc.* 56, 1575–1577. doi: 10.1111/j.1532-5415.2008.01777.x

Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 155–159. doi: 10.1037/0033-2909.112.1.155

Dadashi, F., Millet, G. P., and Aminian, K. (2014). Estimation of front-crawl energy expenditure using wearable inertial measurement units. *IEEE Sens. J.* 14, 1020–1027. doi: 10.1109/JSEN.2013.2292585

Del Din, S., Godfrey, A., Galna, B., Lord, S., and Rochester, L. (2016). Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length. *J. Neuroeng. Rehabil.* 13:46. doi: 10.1186/s12984-016-0154-5

Del Din, S., Kirk, C., Yarnall, A. J., Rochester, L., and Hausdorff, J. M. (2021). Body-worn sensors for remote monitoring of Parkinson's disease motor symptoms: vision, state of the art, and challenges ahead. *J. Parkinsons. Dis.* 11, S35–S47. doi: 10.3233/JPD-202471

Delbaere, K., Close, J. C. T., Mikolaizak, A. S., Sachdev, P. S., Brodaty, H., and Lord, S. R. (2010). The falls efficacy scale international (FES-I). A comprehensive longitudinal validation study. *Age Ageing* 39, 210–216. doi: 10.1093/ageing/afp225

Delbaere, K., Crombez, G., Vanderstraeten, G., Willems, T., and Cambier, D. (2004). Fear-related avoidance of activities, falls and physical frailty. A prospective community-based cohort study. *Age Ageing* 33, 368–373. doi: 10.1093/ageing/afh106

Doheny, E. P., Fan, C. W., Foran, T., Greene, B. R., Cunningham, C., and Kenny, R. A. (2011). "An instrumented sit-to-stand test used to examine differences between older fallers and non-fallers," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, (Boston, MA: IEEE), 3063–3066. doi: 10.1109/IEMBS.2011.6090837

Doheny, E. P., Walsh, C., Foran, T., Greene, B. R., Fan, C. W., Cunningham, C., et al. (2013). Falls classification using tri-axial accelerometers during the five-times-sit-to-stand test. *Gait Posture* 38, 1021–1025. doi: 10.1016/j.gaitpost.2013.05.013

Duncan, R. P., Leddy, A. L., and Earhart, G. M. (2011). Five times sit-to-stand test performance in Parkinson's disease. *Arch. Phys. Med. Rehabil.* 92, 1431–1436. doi: 10.1016/j.apmr.2011.04.008

El-Gohary, M., Pearson, S., McNames, J., Mancini, M., Horak, F., Mellone, S., et al. (2014). Continuous monitoring of turning in patients with movement disability. *Sensors* 14, 356–369. doi: 10.3390/s140100356

Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., et al. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov. Disord.* 22, 1689–1707. doi: 10.1002/mds.21507

Evers, L. J. W., Raykov, Y. P., Krijthe, J. H., Silva De Lima, A. L., Badawy, R., Claes, K., et al. (2020). Real-life gait performance as a digital biomarker for motor fluctuations: the Parkinson@Home validation study. *J. Med. Internet Res.* 22:e19068. doi: 10.2196/19068

Feltz, D. L., and Payment, C. A. (2005). Self-efficacy beliefs related to movement and mobility. *Quest* 57, 24–36. doi: 10.1080/00336297.2005.10491840

Frazier, L. D. (2000). Coping with disease-related stressors in Parkinson's disease. *Gerontologist* 40, 53–63. doi: 10.1093/geront/40.1.53

Galperin, I., Hillel, I., Del Din, S., Bekkers, E. M. J., Nieuwboer, A., Abbruzzese, G., et al. (2019). Associations between daily-living physical activity and laboratory-based assessments of motor severity in patients with falls and Parkinson's disease. *Parkinsonism Relat. Disord.* 62, 85–90. doi: 10.1016/j.parkreldis.2019.01.022

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., et al. (2008). Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov. Disord.* 23, 2129–2170. doi: 10.1002/mds.22340

Goldberg, A., Chavis, M., Watkins, J., and Wilson, T. (2012). The five-times-sit-to-stand test: validity, reliability and detectable change in older females. *Aging Clin. Exp. Res.* 24, 339–344. doi: 10.1007/BF03325265

Greene, B. R., Odonovan, A., Romero-Ortuno, R., Cogan, L., Scanaill, C. N., and Kenny, R. A. (2010). Quantitative falls risk assessment using the timed up and go test. *IEEE Trans. Biomed. Eng.* 57, 2918–2926. doi: 10.1109/TBME.2010.2083659

Haertner, L., Elshehabi, M., Zaunbrecher, L., Pham, M. H., Maetzler, C., van Uem, J. M. T., et al. (2018). Effect of fear of falling on turning performance in Parkinson's disease in the lab and at home. *Front. Aging Neurosci.* 10:78. doi: 10.3389/fnagi.2018.00078

Hoehn, M. M., and Yahr, M. D. (2001). Parkinsonism: onset, progression, and mortality. 1967. *Neurology* 57, S11–S26.

Howcroft, J., Kofman, J., and Lemaire, E. D. (2013). Review of fall risk assessment in geriatric populations using inertial sensors. *J. Neuroeng. Rehabil.* 10:91. doi: 10.1186/1743-0003-10-91

Ivarsson, A., Andersen, M. B., Johnson, U., and Lindwall, M. (2013). To adjust or not adjust: nonparametric effect sizes, confidence intervals, and real-world meaning. *Psychol. Sport Exerc.* 14, 97–102. doi: 10.1016/j.psychsport.2012.07.007

Jonasson, S. B., Nilsson, M. H., Lexell, J., and Carlsson, G. (2018). Experiences of fear of falling in persons with Parkinson's disease - a qualitative study. *BMC Geriatr.* 18:44. doi: 10.1186/s12877-018-0735-1

Kaspar, R., Oswald, F., Wahl, H. W., Voss, E., and Wettstein, M. (2015). Daily mood and out-of-home mobility in older adults: does cognitive impairment matter? *J. Appl. Gerontol.* 34, 26–47. doi: 10.1177/0733464812466290

Latt, M. D., Menz, H. B., Fung, V. S., and Lord, S. R. (2009). Acceleration patterns of the head and pelvis during gait in older people with Parkinson's disease: a comparison of fallers and nonfallers. *J. Gerontol. A Biol. Sci. Med. Sci.* 64, 700–706. doi: 10.1093/gerona/glp009

Lindholm, B., Hagell, P., Hansson, O., and Nilsson, M. H. (2015). Prediction of falls and/or near falls in people with mild Parkinson's disease. *PLoS One* 10:e0117018. doi: 10.1371/journal.pone.0117018

Maetzler, W., Rochester, L., Bhidayasiri, R., Espay, A. J., Sánchez-Ferro, A., and van Uem, J. M. T. (2021). Modernizing daily function assessment in Parkinson's disease using capacity, perception, and performance measures. *Mov. Disord.* 36, 76–82. doi: 10.1002/mds.28377

Marschollek, D., Nemitz, G., Gietzelt, M., Wolf, K. H., Meyer zu Schwabedissen, H., and Haux, R. (2009). Prädiktion von stürzen stationärer patienten in einer geriatrischen klinik: eine klinische studie zur kombination von assessmentdaten und einfachen, sensorisch erfassten gangparametern. *Z. Gerontol. Geriatr.* 42, 317–321. doi: 10.1007/s00391-009-0035-7

Millor, N., Lecumberri, P., Gomez, M., Martinez, A., Martinikorena, J., Rodriguez-Manas, L., et al. (2017). Gait velocity and chair sit-stand-sit performance improves current frailty-status identification. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 2018–2025. doi: 10.1109/TNSRE.2017.2699124

Mukaka, M. M. (2012). Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.* 24, 69–71.

Najafi, B., Aminian, K., Loew, F., Blanc, Y., and Robert, P. A. (2002). Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly. *IEEE Trans. Biomed. Eng.* 49, 843–851. doi: 10.1109/TBME.2002.800763

Narayanan, M. R., Scalzi, M. E., Redmond, S. J., Lord, S. R., Celler, B. G., and Lovell, N. H. (2008). "A wearable triaxial accelerometry system for longitudinal assessment of falls risk," in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'08 - "Personalized Healthcare through Technology"*, (New Jersey, NY: IEEE), 2840–2843. doi: 10.1109/iembs.2008.4649794

Owsley, C., and McGwin, G. (2004). Association between visual attention and mobility in older adults. *J. Am. Geriatr. Soc.* 52, 1901–1906. doi: 10.1111/j.1532-5415.2004.52516.x

Pourghayoomi, E., Behzadipour, S., Ramezani, M., Joghataei, M. T., and Shahidi, G. A. (2020). A new postural stability-indicator to predict the level of fear of falling in Parkinson's disease patients. *Biomed. Eng. Online* 19:64. doi: 10.1186/s12938-020-00808-w

Qiu, H., Rehman, R. Z. U., Yu, X., and Xiong, S. (2018). Application of wearable inertial sensors and a new test battery for distinguishing retrospective fallers from non-fallers among community-dwelling older people. *Sci. Rep.* 8:16349. doi: 10.1038/s41598-018-34671-6

Rai, S. N., and Singh, P. (2020). Advancement in the modelling and therapeutics of Parkinson's disease. *J. Chem. Neuroanat.* 104:101752. doi: 10.1016/j.jchemneu.2020.101752

Regterschot, G. R. H., Folkersma, M., Zhang, W., Baldus, H., Stevens, M., and Zijlstra, W. (2014). Sensitivity of sensor-based sit-to-stand peak power to the effects of training leg strength, leg power and balance in older adults. *Gait Posture* 39, 303–307. doi: 10.1016/j.gaitpost.2013.07.122

Rehman, R. Z. U., Del Din, S., Guan, Y., Yarnall, A. J., Shi, J. Q., and Rochester, L. (2019). Selecting clinically relevant gait characteristics for classification of

early Parkinson's disease: a comprehensive machine learning approach. *Sci. Rep.* 9:17269. doi: 10.1038/s41598-019-53656-7

Rochat, S., Büla, C. J., Martin, E., Seematter-Bagnoud, L., Karmaniola, A., Aminian, K., et al. (2010). What is the relationship between fear of falling and gait in well-functioning older persons aged 65 to 70 years? *Arch. Phys. Med. Rehabil.* 91, 879–884. doi: 10.1016/j.apmr.2010.03.005

Rudman, D. L., Friedland, J., Chipman, M., and Sciortino, P. (2006). Holding on and letting go: the perspectives of pre-seniors and seniors on driving self-regulation in later life. *Can. J. Aging* 25, 65–76. doi: 10.1353/cja.2006.0031

Salarian, A., Horak, F. B., Zampieri, C., Carlson-Kuhta, P., Nutt, J. G., and Aminian, K. (2010). ITUG, a sensitive and reliable measure of mobility. *IEEE Trans. Neural Syst. Rehabil. Eng.* 18, 303–310. doi: 10.1109/TNSRE.2010.2047606

Salkovic, D., Hobert, M. A., Bellut, C., Funer, F., Renno, S., Haertner, L., et al. (2017). Evidence for a selectively regulated prioritization shift depending on walking situations in older adults. *Front. Aging Neurosci.* 9:75. doi: 10.3389/fnagi.2017.00075

Shah, V. V., McNames, J., Mancini, M., Carlson-Kuhta, P., Spain, R. I., Nutt, J. G., et al. (2020b). Laboratory versus daily life gait characteristics in patients with multiple sclerosis, Parkinson's disease, and matched controls. *J. Neuroeng. Rehabil.* 17:159. doi: 10.1186/s12984-020-00781-4

Shah, V. V., McNames, J., Mancini, M., Carlson-Kuhta, P., Nutt, J. G., El-Gohary, M., et al. (2020a). Digital biomarkers of mobility in Parkinson's disease during daily living. *J. Parkinsons. Dis.* 10, 1099–1111. doi: 10.3233/JPD-201914

Shah, V. V., McNames, J., Harker, G., Mancini, M., Carlson-Kuhta, P., Nutt, J. G., et al. (2020c). Effect of bout length on gait measures in people with and without parkinson's disease during daily life. *Sensors* 20:5769.

Staartjes, V. E., and Schröder, M. L. (2018). The five-repetition sit-to-stand test: evaluation of a simple and objective tool for the assessment of degenerative pathologies of the lumbar spine. *J. Neurosurg. Spine* 29, 380–387. doi: 10.3171/2018.2.SPINE171416

Storm, F. A., Nair, K. P. S., Clarke, A. J., Van der Meulen, J. M., and Mazzà, C. (2018). Free-living and laboratory gait characteristics in patients with multiple sclerosis. *PLoS One* 13:e0196463. doi: 10.1371/journal.pone.0196463

Takayanagi, N., Sudo, M., Yamashiro, Y., Lee, S., Kobayashi, Y., Niki, Y., et al. (2019). Relationship between daily and in-laboratory gait speed among healthy community-dwelling older adults. *Sci. Rep.* 9:3496. doi: 10.1038/s41598-019-39695-0

Toosizadeh, N., Mohler, J., Lei, H., Parvaneh, S., Sherman, S., and Najafi, B. (2015). Motor performance assessment in Parkinson's disease: association between objective in-clinic, objective in-home, and subjective/semi-objective measures. *PLoS One* 10:e0124763. doi: 10.1371/journal.pone.0124763

Van Ancum, J. M., van Schooten, K. S., Jonkman, N. H., Huijben, B., van Lummel, R. C., Meskers, C. G. M., et al. (2019). Gait speed assessed by a 4-m walk test is not representative of daily-life gait speed in community-dwelling adults. *Maturitas* 121, 28–34. doi: 10.1016/j.maturitas.2018.12.008

Van Lummel, R. C., Walgaard, S., Maier, A. B., Ainsworth, E., Beek, P. J., and Van Dieën, J. H. (2016). The instrumented Sit-To-Stand test (iSTS) has greater clinical relevance than the manually recorded sit-to-stand test in older adults. *PLoS One* 11:e0157968. doi: 10.1371/journal.pone.0157968

Wang, C., Patriquin, M., Vaziri, A., and Najafi, B. (2021). Mobility performance in community-dwelling older adults: potential digital biomarkers of concern about falling. *Gerontology* 67, 365–373. doi: 10.1159/000512977

Warmerdam, E., Hausdorff, J. M., Atrsaei, A., Zhou, Y., Mirelman, A., Aminian, K., et al. (2020). Long-term unsupervised mobility assessment in movement disorders. *Lancet Neurol.* 19, 462–470. doi: 10.1016/S1474-4422(19)30397-7

Weiss, A., Brozgol, M., Dorfman, M., Herman, T., Shema, S., Giladi, N., et al. (2013). Does the evaluation of gait quality during daily life provide insight into fall risk? A novel approach using 3-Day accelerometer recordings. *Neurorehabil. Neural Repair.* 27, 742–752. doi: 10.1177/1545968313491004

Weiss, A., Herman, T., Plotnik, M., Brozgol, M., Giladi, N., and Hausdorff, J. M. (2011). An instrumented timed up and go: the added value of an accelerometer for identifying fall risk in idiopathic fallers. *Physiol. Meas.* 32, 2003–2018. doi: 10.1088/0967-3334/32/12/009

Williams, J. M., and Nyman, S. R. (2018). Association between the instrumented timed up and go test and cognitive function, fear of falling and quality of life in community dwelling people with dementia. *J. Frailty Sarcopenia Falls* 03, 185–193. doi: 10.22540/jfsf-03-185

Yardley, L., Beyer, N., Hauer, K., Kempen, G., Piot-Ziegler, C., and Todd, C. (2005). Development and initial validation of the falls efficacy scale-international (FES-I). *Age Ageing* 34, 614–619. doi: 10.1093/ageing/afi196

Zhang, W., Regterschot, G. R. H., Geraedts, H., Baldus, H., and Zijlstra, W. (2017). Chair rise peak power in daily life measured with a pendant sensor associates with mobility, limitation in activities, and frailty in old people. *IEEE J. Biomed. Heal. Informatics* 21, 211–217. doi: 10.1109/JBHI.2015.2501828

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

SYSTEMATIC REVIEW
published: 18 January 2022
doi: 10.3389/fnagi.2021.764826

Which Exercise Interventions Can Most Effectively Improve Reactive Balance in Older Adults? A Systematic Review and Network Meta-Analysis

Youngwook Kim[1], Michael N. Vakula[1], David A. E. Bolton[1], Christopher J. Dakin[1], Brennan J. Thompson[1], Timothy A. Slocum[2], Masaru Teramoto[3] and Eadric Bressel[1]*

[1] Department of Kinesiology and Health Science, Utah State University, Logan, UT, United States, [2] Department of Special Education and Rehabilitation Counseling, Utah State University, Logan, UT, United States, [3] Division of Physical Medicine & Rehabilitation, University of Utah, Salt Lake City, UT, United States

OPEN ACCESS

Edited by:
Ondrej Bezdicek, Charles University, Czechia

Reviewed by:
Khaled Trabelsi, University of Sfax, Tunisia
Yoshiro Okubo, Neuroscience Research Australia, Australia

****Correspondence:***
Youngwook Kim
youngwook.kim@usu.edu

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *26 August 2021*
Accepted: *24 December 2021*
Published: *18 January 2022*

Citation:
Kim Y, Vakula MN, Bolton DAE, Dakin CJ, Thompson BJ, Slocum TA, Teramoto M and Bressel E (2022) Which Exercise Interventions Can Most Effectively Improve Reactive Balance in Older Adults? A Systematic Review and Network Meta-Analysis. Front. Aging Neurosci. 13:764826. doi: 10.3389/fnagi.2021.764826

Background: Reactive balance is the last line of defense to prevent a fall when the body loses stability, and beneficial effects of various exercise-based interventions on reactive balance in older adults have been reported. However, their pooled evidence on the relative effects has yet to be described.

Objective: To review and evaluate the comparative effectiveness of various exercise-based interventions on reactive balance in older adults.

Methods: Nine electronic databases and reference lists were searched from inception to August 2021. Eligibility criteria according to PICOS criteria were as follows: (1) population: older adults with the mean age of 65 years or above; (2) intervention and comparison: at least two distinct exercise interventions or one exercise intervention with a no-exercise controlled intervention (NE) compared in each trial; (3) outcome: at least one measure of reactive balance; (4) study: randomized controlled trial. The main network meta-analysis was performed on data from the entire older adult population, involving all clinical conditions as well as healthy older adults. Subgroup analyses stratified by characteristics of participants (healthy only) and reactive balance outcomes (simulated slip or trip while walking, simulated forward falls, being pushed or pulled, and movable platform) were also conducted.

Results: Thirty-nine RCTs (n = 1388) investigating 17 different types of exercise interventions were included in the network meta-analysis. Reactive balance training as a single intervention presented the highest probability (surface under the cumulative ranking (SUCRA) score) of being the best intervention for improving reactive balance and the greatest relative effects vs. NE in the entire sample involving all clinical conditions [SUCRA = 0.9; mean difference (95% Credible Interval): 2.7 (1.0 to 4.3)]. The results were not affected by characteristics of participants (i.e., healthy older adults only) or reactive balance outcomes.

Summary/Conclusion: The findings from the NMA suggest that a task-specific reactive balance exercise could be the optimal intervention for improving reactive balance in older adults, and power training can be considered as a secondary training exercise.

Keywords: older adults, aging, balance, reactive balance, exercise, falls, accidental falls, fall prevention

INTRODUCTION

The World Health Organization reported that approximately 28–35% of people aged 65 or above experience at least one fall each year, and the frequency of falls increases with age and frailty level (World Health Organization, 2008). Among various intrinsic risk factors for falls, gait and balance problems have been considered the strongest risk factors (Deandrea et al., 2010; Ambrose et al., 2013). Balance can be mechanistically achieved and maintained by a complex set of sensorimotor control systems including the multisensory (visual, somatosensory, and vestibular system) integration into the central nervous system and the subsequent motor output of the musculoskeletal system (Shumway-Cook and Woollacott, 2017). However, older adults show age-related decline in sensorimotor systems, which in turn increases the risks of falls (Mahoney et al., 2019; Osoba et al., 2019). Given the inherent and inevitable age-related degeneration in sensorimotor systems, it is becoming increasingly clear that in order to prevent potential repercussions, such as aging-related disease, disabilities, injuries, and falls, there is an urgent need for effective interventions to decelerate or even reverse the retrogression in the balance and gait control systems (Kim et al., 2020; Sibley et al., 2021).

In daily life, reactive balance, referred to as the ability to control balance in response to mechanical disturbances, plays a critical role in avoiding and adapting to the complex environments that menace postural stability. The WHO Global Report on Falls Prevention in Older Age reported that factors related to the physical environment, for instance, uneven sidewalks, unmarked obstacles, and slippery surfaces, are some of the most common causes (30–50%) of falls in older adults (World Health Organization, 2008). Notably, slips and trips were the most prevalent causes of falls in regards to circumstances in older adults (Berg et al., 1997). Reactive balance strategies, such as swaying around the ankle or hip joints, taking a reactive step, or reaching to grasp a handhold (Shumway-Cook and Woollacott, 2017), need to be executed promptly so as to avoid falls following a postural perturbation. In the same vein, the balance recovery reactions have also shown age-related differences in older adults vs. young adults and in fallers vs. non-fallers (Alissa et al., 2020; Okubo et al., 2021).

There is a considerable amount of literature on the effects of a variety of interventions on reactive balance, including several systematic reviews and meta-analyses focusing on older adults (Bohm et al., 2015; Lesinski et al., 2015; McCrum et al., 2017; Moore et al., 2019). However, there remain some limitations in the prior syntheses. First, the exercise interventions were limited to only those interventions focused on balance or strength training despite the existence of many studies that use exercises that specifically train reactive balance. Consequently, to the best of our knowledge, none of the previous reviews or meta-analyses have considered comparative effects between different types of exercises and the efficacy of multifaceted exercise interventions with more than one type of exercise on reactive balance. Thus, there is a need for a more comprehensive and inclusive analysis utilizing precise coding of exercise types targeting specific biological systems and functional aspects for better prescriptive guidance (Sibley et al., 2021). Second, the systematic review by Moore et al. (2019) who examined the effectiveness of active physical training interventions on reactive balance did not perform a quantitative synthesis (Moore et al., 2019). Consequently, there remains a lack of pooled evidence on the relative effects of different exercise interventions on reactive balance. Moreover, a conventional pairwise meta-analysis is restricted to a head-to-head comparison of only two different interventions, and thus, RCTs with other types of exercise interventions, that are also effective, can potentially be excluded. To tackle this problem, a network meta-analysis (NMA) is well suited, because it facilitates comparisons of multiple pairs of interventions in one statistical model (Dias et al., 2018).

Therefore, the current study aimed to quantitatively synthesize the available evidence of RCTs in detail using a systematic review and NMA to: (1) combine information from all available randomized comparisons of a set of exercise interventions for reactive balance in older adults; (2) to appraise the relative effects of different exercise interventions on reactive balance; and (3) to determine the ranking of each to provide practical and clinical suggestions to design evidence-based exercise programs for reactive balance. The research question was as follows: "What type of exercise intervention is most effective in improving overall measures as well as each measure of reactive balance in older adults?".

METHODS

The protocol was prospectively registered in the PROSPERO database (CRD42021256638) and conducted in accordance with the PRISMA extension statement for network meta-analysis (Hutton et al., 2015).

Eligibility Criteria

The population of interest included older adults with a sample mean age of 65 years or above with no restriction on the injury or disorder type and with no history of falls studied in various research settings (e.g., community, clinics, and long-term care facilities). Studies were included, if at least two experimental groups participated in each of the different exercise interventions or if there was at least one exercise intervention group

TABLE 1 | Definitions of exercise types.

Exercise type	Code	Definitions
Single balance exercise including reactive balance component	SBR	An intervention including a balance exercise with one or more mechanical postural perturbations given during the exercise
Single balance exercise not including reactive balance component	SBNR	An intervention including a balance exercise without any mechanical postural perturbations
Multiple balance exercises including reactive balance component	MBR	An intervention including more than one type of balance exercise with one or more mechanical postural perturbations given during one of the exercises
Multiple balance exercises not including reactive balance component	MBNR	An intervention including more than one type of balance exercise without any mechanical postural perturbations
Unspecified balance exercise	balUS	Balance exercise without any details given in the original article
Gait training including reactive balance component	gaitR	An intervention including gait training with one or more mechanical postural perturbations given during the exercise
Gait training not including reactive balance component	gaitNR	An intervention including gait training without any mechanical postural perturbations
Whole body vibration	WBV	Any activity performed on a machine with a vibrating platform
Strength	str	Exercise that uses the external resistance load (e.g., body weight, resistance bands, machines) to force skeletal muscles contract.
Power	pw	Exercise that applies the maximum amount of force (muscle contraction against a resistance) in the shortest period of time.
3D exercise	3d	Exercise that requires multi-dimensional movements with a specific name of the exercise (e.g., Yoga, dance, Tai Chi)
Flexibility	flex	Exercise that intends to restore or maintain the optimal range of motion (ROM) available to a joint or joints.
Functional training	FT	Exercise that utilizes functional activities as the training stimulus that is based on the theoretical concept of task specificity
Aerobic	aer	Exercise aimed at cardiovascular conditioning. It is aerobic in nature and simultaneously increases the heart rate and the return of blood to the heart.
No exercise	NE	A group received none of the exercise interventions listed above

with a no-exercise control group. Studies involving any non-exercise interventions (e.g., medication, electrical stimulation, or nutritional supplement) were excluded. Details regarding the exercise interventions must have been provided in the reports. The studies must have included at least one reactive balance assessment, which is defined in this study as an assessment using mechanical postural perturbation during a static or dynamic steady-state task. The studies included in this review were restricted to randomized controlled trials (RCTs) and written in the English language.

Search Strategy

The following electronic databases were initially searched by one reviewer (Y.K.) from the inception to February 2021: MEDLINE, EBSCO, CINAHL, SPORTDiscus, PsycINFO, PubMed. WorldCat.org, OpenGrey.eu, and PROQUEST were additionally searched for unpublished trials. To keep this search up to date, an updated search followed in August 2021 by two reviewers (YK and MV). Earlier reviews and bibliographies of included studies were reviewed for additional potentially relevant trials. The combination of the following keywords was employed for the database searches: (aged OR aging OR old* OR elder* OR senior*) AND (exercise OR train* OR activit* OR rehabilitat* OR therap* OR physiotherapy OR hydrotherapy OR conditioning OR exertion OR recreation* OR aerobic* OR stretch* OR strengthen* OR walk* OR jog* OR run* OR cycl* OR pilates OR yoga OR tai chi OR ai chi OR dance OR swim*) AND (reactive postural response OR stepping response OR perturbation OR slip perturbation OR reactive balance OR reactive stepping OR protective stepping OR compensatory stepping OR anticipatory postural adjustment* OR compensatory postural adjustment* OR anticipatory postural response* OR compensatory postural response* OR anticipatory adjustment* OR compensatory adjustment* OR postural adaptation* OR postural stabili*ation OR automatic postural response* OR postural stepping response*) AND (random*).

Study Selection

After exporting the references and removing duplicates, titles and abstracts of records were screened independently by two reviewers (YK and MV) according to the eligibility criteria. Full texts of all potentially relevant trials were subsequently retrieved and reviewed to confirm the final eligible trials. Any disagreements were resolved via consensus, and when any disagreement was elusive, a third reviewer (EB) acted as an arbiter.

Data Extraction and Coding

A total of 46 eligible studies were reviewed and coded in REDCap (https://www.projectredcap.org/) by one reviewer (YK) and confirmed by a second reviewer (M.V.). Any disagreements were resolved via consultation with a third reviewer (EB). The

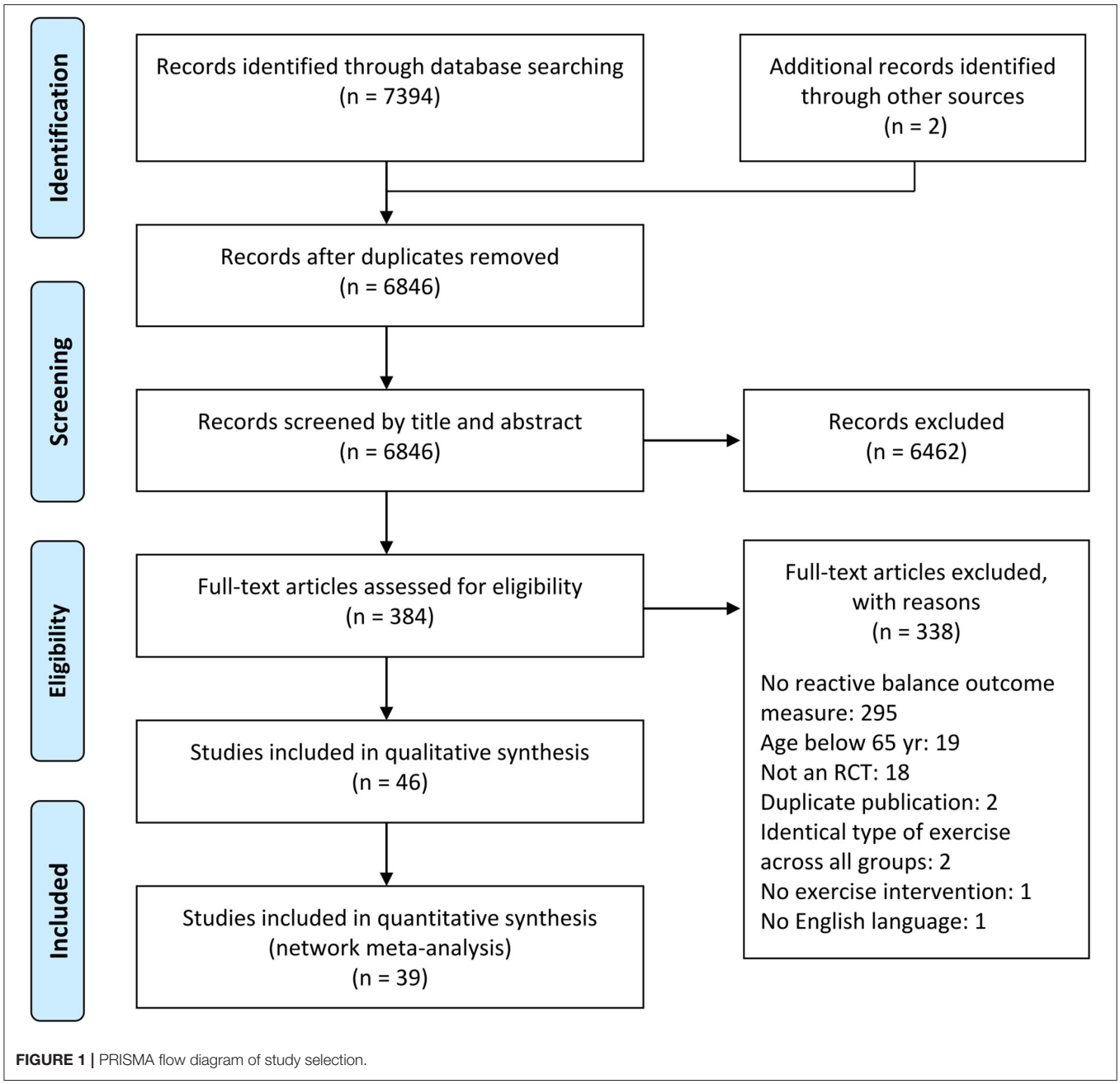


FIGURE 1 | PRISMA flow diagram of study selection.

extracted data included: (1) study characteristics; (2) baseline demographics of participants; (3) exercise interventions; (4) reactive balance outcome measures; and (5) results. Exercise categorizations developed by Howe et al. and Sibley et al. were modified in consideration of the purpose of the current research and applied to the coding (**Table 1**) (Howe et al., 2011; Sibley et al., 2021).

Means (M) and standard deviations (SD) for all eligible outcomes of reactive balance measures at baseline and post-intervention were extracted for the analysis. Missing data related to eligibility and study outcomes (i.e., data not reported either in a text or on publicly accessible data repositories) were requested to the corresponding authors via email. In the case of no response after one month, a second request was sent, if another month elapsed without response, the data were considered irretrievable. If the requested, but not retrieved data were presented in a graphical format rather than numeric data (e.g., tabular format), Engauge Digitizer 12.1 (http://markummitchell.github.io/engauge-digitizer/) was applied for data digitization and extraction.

Risk of Bias

To ascertain an overall and study-level risk of bias for each trial, a pair of reviewers (YK and MV) independently determined the

TABLE 2 | Characteristics of participants.

Study	Disease category	Sample size (post-intervention)	Attrition rate (%)	Age (years)
Allin et al. (2020)	Healthy	34 (29)	15	70.4
Arampatzis et al. (2011)	Healthy	55 (38)	31	67.7
Arghavani et al. (2020)	Healthy (fallers: 6 months)	60 (49)	18	69.6
Beling and Roller (2009)	Healthy	23 (19)	17	80.0
Bieryla et al. (2007)	Healthy	12 (11)	8	73.3
Bogaerts et al. (2007)	Healthy	220 (161)	27	67.1
Cabrera-Martos et al. (2020)	Parkinson's	44 (44)	0	76.5
Cherup et al. (2019)	Parkinson's	42 (35)	17	71.2
Chyu et al. (2010)	Postmenopausal women with osteopenia	61 (53)	13	71.9
Donath et al. (2016)	Healthy	59 (48)	19	69.7
Gatts and Woollacott (2007)	Healthy (balance deficiency without any neurological disorder); Arthritis, back, knee, or hip surgery not excluded.	22 (19)	14	77.6
Gatts (2008)	Healthy (balance deficiency without any neurological disorder); Arthritis, back, knee, or hip surgery not excluded.	22 (19)	14	77.6
Granacher et al. (2006)	Healthy	60 (60)	0	66.5
Granacher et al. (2009)	Healthy	40 (40)	0	67.0
Hamed et al. (2018)	Healthy	63 (47)	25	71.2
Hatzitaki et al. (2009)	Healthy	56 (56)	0	70.9
Hu and Woollacott (1994)	Healthy	24 (24)	0	75.2
Inacio et al. (2018)	Healthy	18 (18)	0	71.9
Jagdhane et al. (2016)	Healthy	6 (6)	0	73.3
Kim and Lockhart (2010)	Healthy	18 (18)	0	NS
Klamroth et al. (2019)	Parkinson's	43 (37)	14	65.3
Lacroix et al. (2016)	Healthy	66 (60)	9	72.8
Li et al. (2009)	Healthy	50 (40)	20	65.3
Ma et al. (2019)	Healthy	33 (24)	27	69.8
Mansfield et al. (2010)	Healthy (fallers: 5 years)	34 (30)	12	69.7
Marigold et al. (2005)	chronic stroke	59 (48)	19	67.8
Morat et al. (2019)	Healthy	51 (45)	12	69.4
Ni et al. (2014)	Healthy	48 (39)	19	74.2
Ochi et al. (2015)	Healthy	20 (20)	0	80.6
Okubo et al. (2019)	Healthy	44 (41)	7	72.1
Pamukoff et al. (2014)	Healthy (some lower extremity mobility dysfunction)	20 (15)	25	70.8
Parijat and Lockhart (2012)	Healthy	24 (24)	0	72.7
Parijat et al. (2015a)	Healthy	24 (24)	0	72.4
Parijat et al. (2015b)	Healthy	24 (24)	0	72.4
Pluchino et al. (2012)	Healthy	40 (27)	33	72.1
Qutubuddin et al. (2007)	Parkinson's	22 (15)	32	72.8
Rieger et al. (2020)	Healthy	30 (30)	0	71.0
Rossi et al. (2014)	Healthy	46 (46)	0	67.5
Santos et al. (2017)	Parkinson's	40 (40)	0	67.8
Schlenstedt et al. (2015)	Parkinson's	40 (32)	20	75.7
Shimada et al. (2003)	Healthy	34 (32)	6	80.9
Sohn and Kim (2015)	Healthy	18 (18)	0	73.7
Thomas and Kalicinski (2016)	Healthy	24 (24)	0	67.1
Wang et al. (2019)	Healthy	146 (146)	0	72.7
Wolf et al. (1997)	Healthy	72 (54)	25	76.9
Wooten et al. (2018)	Healthy (fallers: 1 year)	30 (16)	47	72.6

TABLE 3 | Summary of exercise interventions.

Study	Dosage		Total duration (week)	Exercise interventions		
	Min/session	Time/week		Group1	Group2	Group3
Allin et al. (2020)	30-60	2	2	SBR + gaitR	MBNR + gaitNR + str	
Arampatzis et al. (2011)	90	2	14	MBR	SBNR + str	NE
Arghavani et al. (2020)	60	3	8	SBR	MBNR + gaitNR + str	NE
Beling and Roller (2009)	60	3	12	MBR + gaitNR + flex + str	NE	
Bieryla et al. (2007)	15	1	1	gaitR	gaitNR	
Bogaerts et al. (2007)	40-90	3	1 year	MBNR + WBV	SBNR + str + flex + aer	NE
Cabrera-Martos et al. (2020)	45	3	8	FT	FT + flex	
Cherup et al. (2019)	60	2	12	pw	Str	
Chyu et al. (2010)	60	3	24	3d	NE	
Donath et al. (2016)	66	2	8	3d	MBNR	NE
Gatts and Woollacott (2007)	90	5	3	3d	SBNR + flex	
Gatts (2008)	90	5	3	3d	SBNR + flex	
Granacher et al. (2006)	60	3	13	str	SBNR	NE
Granacher et al. (2009)	60	3	13	str	NE	
Hamed et al. (2018)	90	2	14	str	SBR	NE
Hatzitaki et al. (2009)	30	3	4	SBNR	SBNR	NE
Hu and Woollacott (1994)	60	10 sessions (total)	15 days (total)	SBNR	NE	
Inacio et al. (2018)	15	3	8	pw	str	
Jagdhane et al. (2016)	60	3	4	SBR	NE	
Kim and Lockhart (2010)	NR	NR	8	str	MBNR	NE
Klamroth et al. (2019)	40	2	8	gaitR	gaitNR	
Lacroix et al. (2016)	45	3	12	MBNR + str + pw	MBNR + str + pw	NE
Li et al. (2009)	60	4 for 6weeks, 7 for 10 weeks	16	3d	NE	
Ma et al. (2019)	60	2	12	3d	NE	
Mansfield et al. (2010)	30	3	6	SBR	SBNR + flex	
Marigold et al. (2005)	60	3	10	MBR + gaitNR	SBNR + flex	
Morat et al. (2019)	40	3	8	SBR	SBNR	NE
Ni et al. (2014)	60	2	12	3d	MBNR	3d
Ochi et al. (2015)	30	3	12	MBNR + WBV	SBNR + str	
Okubo et al. (2019)	40	3	1	gaitR	gaitNR	
Pamukoff et al. (2014)	60	3	6	pw	str	
Parijat and Lockhart (2012)	40	1	1	gaitR	gaitNR	
Parijat et al. (2015a)	35–55	1	1	gaitR	gaitNR	
Parijat et al. (2015b)	35–55	1	1	gaitR	gaitNR	
Pluchino et al. (2012)	60	2	8	MBNR + gaitNR	3d	MBNR
Qutubuddin et al. (2007)	30	2	4	balUS	MBNR + gaitNR	
Rieger et al. (2020)	NS	1	1	gaitR	gaitNR	
Rossi et al. (2014)	40	3	6	SBNR	NE	
Santos et al. (2017)	60	2	8	str + flex	MBR + gaitNR	
Schlenstedt et al. (2015)	60	2	7	str	MBR	
Shimada et al. (2003)	40	2–3	12	MBNR	gaitNR	str + flex
Sohn and Kim (2015)	60	3	8	str	balUS	NE
Thomas and Kalicinski (2016)	70	2	6	MBNR	NE	
Wang et al. (2019)	30	1	1	gaitR	gaitNR	
Wolf et al. (1997)	60	1–2	15	MBR	NE	3d
Wooten et al. (2018)	45	3	6	MBNR	3d	

SBR, Single balance exercise including reactive balance component; SBNR, Single balance exercise not including reactive balance component; MBR, Multiple balance exercises including reactive balance component; MBNR, Multiple balance exercises not including reactive balance component; balUS, Unspecified balance exercise; gaitR, Gait training including reactive balance component; gaitNR, Gait training not including reactive balance component; WBV, Whole body vibration; str, Strength; pw, Power; 3d, 3D exercise; FT, Functional training; flex, Flexibility; aer, Aerobic; NE, No exercise.

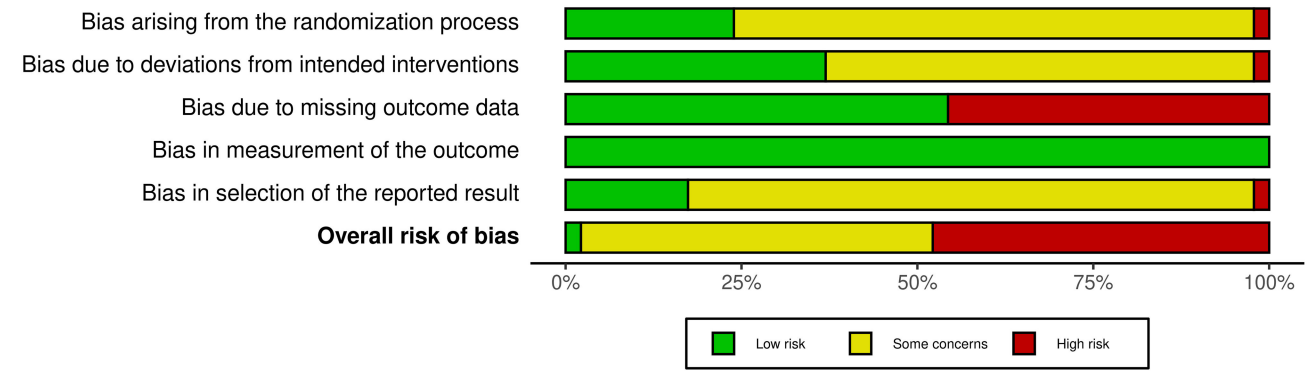


FIGURE 2 | Summary of the distributions of the reviewers' judgements across the studies for each risk of bias domain.

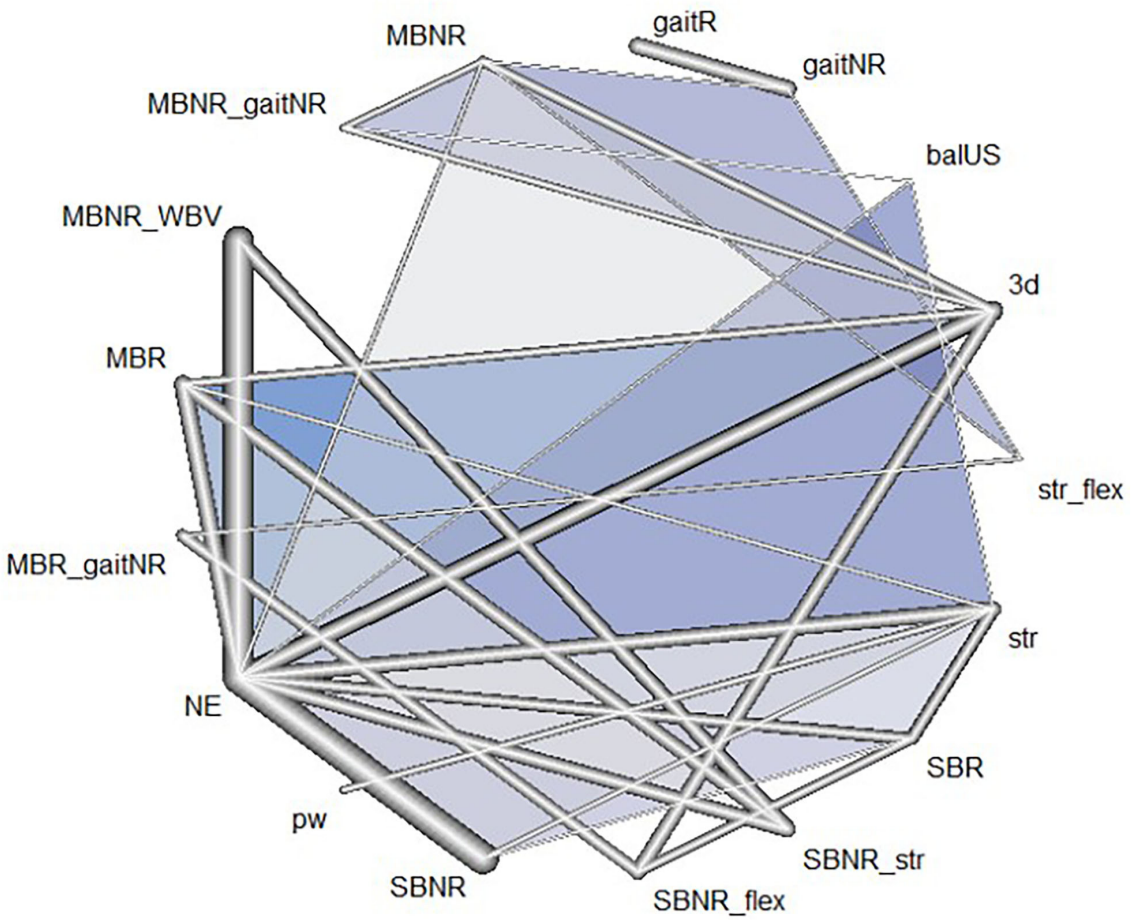


FIGURE 3 | Network geometry of the included exercise programs: Each line indicates a direct comparison of two different exercise programs. The thickness of the edge is proportional to the number of direct comparisons in the network. Different exercise types combined in one program are connected via underscores. The blue triangles refer to multi-arm trials comprised of three exercise programs in the nodes. SBR, Single balance exercise including reactive balance component; SBNR, Single balance exercise not including reactive balance component; MBR, Multiple balance exercises including reactive balance component; MBNR, Multiple balance exercises not including reactive balance component; balUS, Unspecified balance exercise; gaitR, Gait training including reactive balance component; gaitNR, Gait training not including reactive balance component; WBV, Whole body vibration; str, Strength; pw, Power; 3d, 3D exercise; flex, Flexibility; aer, Aerobic; NE, No exercise.

3d															
-0.50 (-3.38, 2.41)	balUS														
-0.07 (-3.80, 3.67)	0.43 (-4.17, 5.04)	gaitNR													
-1.07 (-5.16, 2.96)	-0.58 (-5.45, 4.26)	-1.01 (-2.63, 0.61)	gaitR												
-0.34 (-1.92, 1.22)	0.16 (-2.91, 3.21)	-0.26 (-3.77, 3.22)	0.74 (-3.08, 4.59)	MBNR											
-0.28 (-3.04, 2.43)	0.21 (-2.67, 3.09)	-0.20 (-4.64, 4.20)	0.79 (-3.89, 5.49)	0.07 (-2.78, 2.85)	MBNR + gaitNR										
-0.07 (-3.26, 3.10)	0.42 (-3.57, 4.39)	0.01 (-4.83, 4.72)	1.01 (-4.02, 5.99)	0.27 (-3.06, 3.62)	0.21 (-3.77, 4.21)	MBNR + WBV									
-0.31 (-2.58, 1.97)	0.20 (-3.10, 3.45)	-0.23 (-4.50, 4.00)	0.78 (-3.75, 5.30)	0.04 (-2.55, 2.60)	-0.02 (-3.38, 3.32)	-0.23 (-3.54, 3.09)	MBR								
-0.93 (-4.38, 2.54)	-0.43 (-4.75, 3.95)	-0.85 (-5.15, 3.43)	0.16 (-4.42, 4.74)	-0.58 (-4.08, 2.89)	-0.65 (-4.90, 3.70)	-0.85 (-5.36, 3.71)	-0.63 (-4.58, 3.40)	MBR + gaitNR							
0.25 (-1.14, 1.63)	0.75 (-2.03, 3.51)	0.32 (-3.54, 4.09)	1.33 (-2.80, 5.44)	0.59 (-1.18, 2.37)	0.53 (-2.28, 3.35)	0.32 (-2.54, 3.18)	0.56 (-1.46, 2.57)	1.18 (-2.37, 4.67)	NE						
-1.28 (-4.27, 1.73)	-0.77 (-4.39, 2.79)	-1.20 (-5.86, 3.44)	-0.19 (-5.09, 4.72)	-0.94 (-4.13, 2.28)	-0.98 (-4.78, 2.79)	-1.19 (-5.14, 2.72)	-0.97 (-4.06, 2.16)	-0.36 (-4.73, 4.02)	-1.52 (-4.27, 1.20)	pw					
-0.37 (-2.44, 1.68)	0.14 (-3.01, 3.22)	-0.30 (-4.42, 3.75)	0.72 (-3.69, 5.11)	-0.03 (-2.35, 2.29)	-0.08 (-3.26, 3.11)	-0.29 (-3.55, 2.95)	-0.06 (-2.54, 2.41)	0.55 (-3.24, 4.32)	-0.61 (-2.16, 0.90)	0.91 (-2.10, 3.91)	SBNR				
-0.90 (-3.41, 1.63)	-0.40 (-4.03, 3.23)	-0.82 (-5.01, 3.35)	0.18 (-4.28, 4.63)	-0.56 (-3.32, 2.21)	-0.62 (-4.18, 3.01)	-0.83 (-4.66, 3.01)	-0.59 (-3.79, 2.61)	0.04 (-3.07, 3.09)	-1.14 (-3.74, 1.43)	0.39 (-3.31, 4.05)	-0.53 (-3.47, 2.40)	SBNR + flex			
-0.06 (-3.13, 2.95)	0.44 (-3.41, 4.26)	0.03 (-4.70, 4.68)	1.04 (-3.96, 5.97)	0.29 (-2.97, 3.53)	0.23 (-3.65, 4.11)	0.01 (-2.87, 2.87)	0.25 (-2.74, 3.16)	0.87 (-3.63, 5.25)	-0.31 (-3.06, 2.42)	1.22 (-2.63, 5.00)	0.31 (-2.82, 3.41)	0.83 (-2.95, 4.54)	SBNR + str		
-2.40 (-4.45, -0.37)	-1.91 (-5.06, 1.26)	-2.33 (-6.40, 1.70)	-1.32 (-5.69, 3.05)	-2.07 (-4.37, 0.26)	-2.12 (-5.30, 1.05)	-2.34 (-5.62, 0.93)	-2.10 (-4.62, 0.43)	-1.48 (-5.09, 2.16)	-2.65 (-4.28, -1.04)	-1.13 (-4.17, 1.88)	-2.04 (-4.13, 0.02)	-1.51 (-4.11, 1.14)	-2.35 (-5.49, 0.85)	SBR	
-0.48 (-2.60, 1.64)	0.03 (-2.93, 2.94)	-0.40 (-4.56, 3.74)	0.61 (-3.85, 5.04)	-0.14 (-2.51, 2.28)	-0.19 (-3.32, 2.96)	-0.39 (-3.70, 2.94)	-0.16 (-2.44, 2.14)	0.46 (-3.40, 4.29)	-0.72 (-2.45, 1.03)	0.80 (-1.29, 2.92)	-0.10 (-2.25, 2.04)	0.43 (-2.59, 3.44)	-0.42 (-3.55, 2.76)	1.93 (-0.22, 4.09)	str
-0.80 (-4.12, 2.54)	-0.29 (-4.52, 3.94)	-0.72 (-4.28, 2.79)	0.31 (-3.61, 4.13)	-0.45 (-3.61, 2.69)	-0.52 (-4.60, 3.61)	-0.72 (-5.15, 3.69)	-0.48 (-4.35, 3.43)	0.13 (-2.96, 3.20)	-1.03 (-4.42, 2.37)	0.49 (-3.85, 4.79)	-0.42 (-4.09, 3.29)	0.12 (-3.41, 3.60)	-0.74 (-5.07, 3.66)	1.62 (-1.98, 5.19)	-0.32 (-4.06, 3.40)	str + flex

FIGURE 4 | Relative effect estimates with 95% credible intervals of all pairs of exercise interventions.

bias arising from the following domains using the Cochrane risk of bias tool (RoB 2): (1) randomization process; (2) deviations from the intended interventions; (3) missing outcome data; (4) measurement of the outcome; and (5) selection of the reported result (Sterne et al., 2019). Each domain was assigned a judgement of "low risk," "some concerns," or "high risk." Disagreements were resolved through discussion or referral to a third reviewer (EB).

Data Synthesis and Statistical Analysis

Considering indeterminate baseline similarities of reactive balance measures in several studies, change values from baseline to post-intervention were calculated or directly extracted from the published data. If there were more than one post-intervention measure (e.g., post-intervention and follow-up), only the data immediately following the termination of the intervention phase was used. SDs for changes from baseline (pre) to post-intervention (post) were calculated using the following formula (Higgins et al., 2019):

$$SD_{change} = \sqrt{SD_{pre}{}^2 + SD_{post}{}^2 - 2{*}Corr{*}SD_{pre}{*}SD_{post}}$$

Corr in the SD_{change} equation is the correlation coefficient describing how similar the pre and post-interventions were across participants. When the correlation coefficient was not reported, it was set as 0.5 (Fu et al., 2008; Bruderer-Hofstetter et al., 2018; Lai et al., 2018; Wu et al., 2021). In the case of a lower score signifying better performance in reactive balance measures (e.g., reaction time), scale directions were adjusted by multiplying −1 to the M_{change} data, which led to a greater effect size indicating an improvement. Missing SDs were imputed from standard errors (SE), 90 or 95% confidence intervals (CI). Using the M_{change} and SD_{chang} data, standardized mean differences (SMD) and standard errors (SE) were calculated.

To include multi-arm trials, two approaches were adopted to avoid a unit-of-analysis error (Rücker et al., 2017; Higgins

TABLE 4 | Ranking of exercise interventions.

Bayesian framework			Frequentist framework		
Ranking	**Exercise**	**SUCRA score**	**Ranking**	**Exercise**	**P-score**
1	SBR	0.90	1	SBR	0.94
2	pw	0.67	2	pw	0.70
3	gaitR	0.62	3	gaitR	0.64
4	SBNR + flex	0.58	4	SBNR + flex	0.61
5	MBR + gaitNR	0.58	5	MBR + gaitNR	0.60
6	str + flex	0.55	6	str + flex	0.57
7	balUS	0.49	7	balUS	0.49
8	str	0.49	8	str	0.49
9	SBNR	0.46	9	SBNR	0.46
10	MBNR	0.46	10	MBNR	0.45
11	MBR	0.45	11	MBR	0.44
12	MBNR + gaitNR	0.44	12	MBNR + gaitNR	0.43
13	MBNR + WBV	0.40	13	MBNR + WBV	0.38
14	SBNR + str	0.40	14	SBNR + str	0.37
15	gaitNR	0.39	15	gaitNR	0.37
16	3d	0.35	16	3d	0.33
17	NE	0.27	17	NE	0.23

SBR, Single balance exercise including reactive balance component; SBNR, Single balance exercise not including reactive balance component; MBR, Multiple balance exercises including reactive balance component; MBNR, Multiple balance exercises not including reactive balance component; balUS, Unspecified balance exercise; gaitR, Gait training including reactive balance component; gaitNR, Gait training not including reactive balance component; WBV, Whole body vibration; str, Strength; pw, Power; 3d, 3D exercise; flex, Flexibility; aer, Aerobic; NE, No exercise.

et al., 2019). First, all relevant experimental intervention groups composed of the same categories of exercises were combined into a single group. This step enabled a single pairwise comparison between a combined group and a comparison group in each study. Second, in the case of heterogeneous exercise types across all intervention groups, we included all relevant comparisons as a series of two-arm comparisons and reflect the fact that comparisons within multi-arm studies are correlated (Schwarzer et al., 2015). Accordingly, adjusted SEs of the two-arm comparisons in each multi-arm study were computed using "netmeta" package in R software. The majority of the eligible trials consisted of multiple outcomes in each trial. When multiple SMDs were estimated in a single study, therefore, a pooled SMD with SE was computed.

To estimate the comparative effectiveness of exercise-based interventions on reactive balance, we implemented NMA, which incorporates both direct (i.e., head-to-head comparison from pairwise meta-analysis) and indirect comparisons (i.e., from network meta-analysis) in one statistical model. A Bayesian framework of NMA was conducted using Markov chain Monte Carlo simulations, and non-informative prior distributions for treatment effects were adopted (Lunn et al., 2000; Dias et al., 2018). A random-effects model was used considering the clinical and methodological between-study heterogeneity (Sutton et al., 2000; Borenstein et al., 2009). The NMA was conducted for all available exercise interventions included in at least two trials. The analyses utilized a burn-in period (50,000 iterations) and a follow-up period (100,000 iterations) to minimize bias of initial values when the chain reached its target distribution (Brooks and Gelman, 1998). The convergence was assessed using the trace plot, density plot, and Brooks-Gelman-Rubin diagnostic statistics (Brooks and Gelman, 1998).

The overall geometry of the network was presented in a network graph. Based on Bayesian posterior rank probabilities, the ranking of exercise interventions was estimated using a hierarchical tool, the surface under the cumulative ranking curve (SUCRA) score, measured on a scale from 0 (theoretically the worst) to 1 (the best). In addition, a network forest plot was produced with the "no exercise (NE)" as a reference intervention. The posterior distribution of the SMDs was reported using the mean differences fto the reference intervention with 95% credible intervals (CrI), which indicate that there is a 95% probability that the unobserved (unknown) effect estimates would fall within the intervals (Hespanhol et al., 2019). If a 95% CrI contains zero (i.e., null effect representing the null hypothesis), the effect can be considered statistically insignificant (Hespanhol et al., 2019). The relative effects with 95% CrI of all pairs of exercise interventions were reported in a matrix. Consistency, which is the most important assumption underlying a NMA and indicates agreement between direct and indirect estimates in the network (Salanti et al., 2014), was checked using the node-splitting analysis. The original intention of the first subgroup analysis was to conduct a network meta-analysis stratified by characteristics of participants (i.e., healthy and disease-specific). However, due to the insufficient number of exercise interventions to establish a network in each disease category (e.g., Parkinson's disease), the first subgroup analysis was performed by the inclusion of studies with healthy older adults only (78% of all studies) who did not have any disease, injury, or disability at the time of the studies. The second subgroup analysis was conducted by grouping the outcome measures by the types of reactive balance outcomes: (1) simulated slip or trip while walking; (2) simulated forward falls; (3) being pushed or pulled; (4) movable platform; and (5) balance test battery. A sensitivity analysis was carried out using a frequentist framework NMA to appraise the robustness of the results. Sources of statistical heterogeneity and small study bias were not explored due to the insufficient number of trials ($k \leq 5$) for each comparison. All data syntheses and statistical analyses were conducted using "Gemtc" (version 1.0-1), "rjags" (version 4-10), and "netmeta" (version 1.4-0) packages in R (Version 4.1.0, R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Study Selection

A total of 7,394 records were retrieved from electronic databases and two additional records were obtained from other sources, of which 384 studies remained after removing duplicates and screening titles and abstracts. Based on the full-text screening, 46 records fulfilled the eligibility criteria and thus were included for qualitative analysis (i.e., systematic review), whereas seven studies were additionally excluded from the quantitative analysis (i.e., network meta-analysis) due to data not being reported

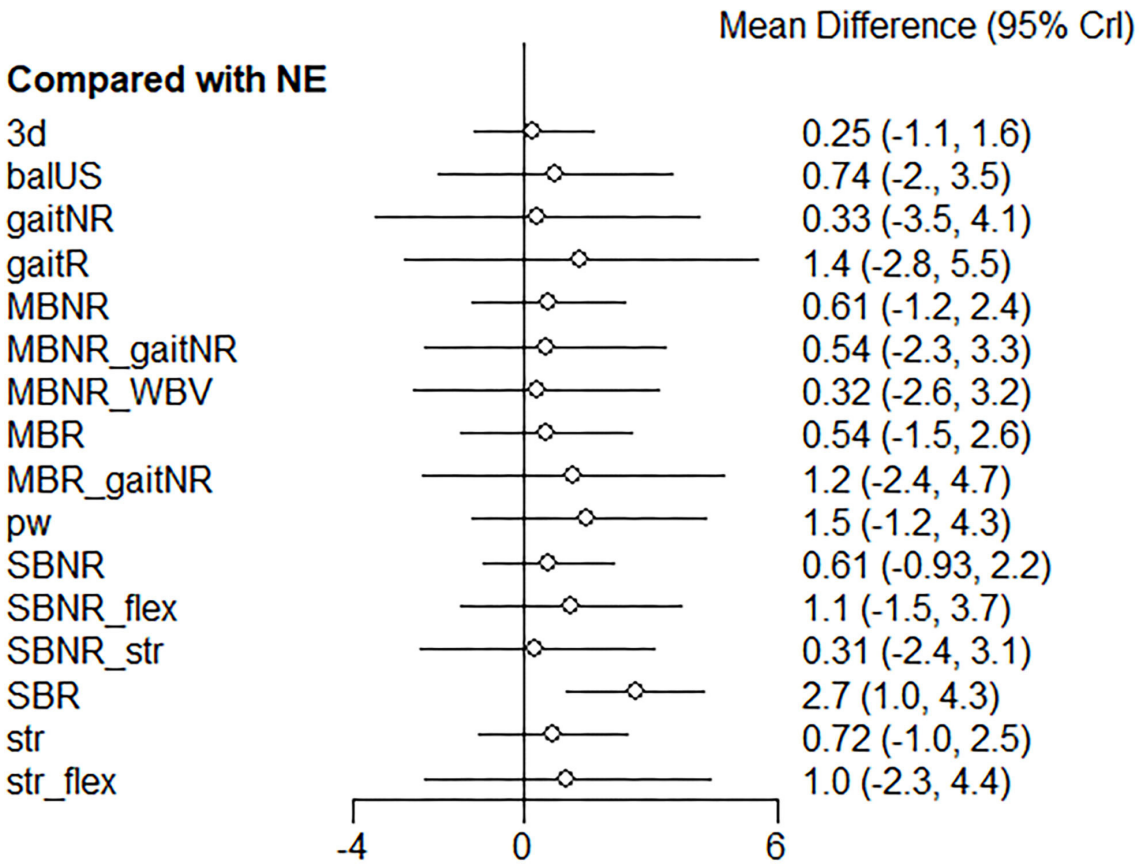


FIGURE 5 | Forest plot of the relative effects of exercise interventions with a no-exercise as a reference group. SBR, Single balance exercise including reactive balance component; SBNR, Single balance exercise not including reactive balance component; MBR, Multiple balance exercises including reactive balance component; MBNR, Multiple balance exercises not including reactive balance component; balUS, Unspecified balance exercise; gaitR, Gait training including reactive balance component; gaitNR, Gait training not including reactive balance component; WBV, Whole body vibration; str, Strength; pw, Power; 3d, 3D exercise; flex, Flexibility; aer, Aerobic; NE, No exercise.

and not irretrievable (Kim and Lockhart, 2010; Okubo et al., 2019; Wang et al., 2019), exercise types not included in the network (Allin et al., 2020; Cabrera-Martos et al., 2020), exercise intervention included in only one trial (Lacroix et al., 2016), and no continuous data reported (Beling and Roller, 2009), resulting in a total of 39 studies for NMA. The schematic flow chart for the selection process is presented in **Figure 1**, and all included studies are listed in **Supplementary Table 1**.

Characteristics of Included Studies

The eligible studies represented a total of 1,745 older adults, included in both pre and post-intervention analyses, with the mean age of 71.9 ± 3.9 years (ranged from 65.3–80.9 years). The majority of the studies exclusively included community-dwelling healthy older adults ($k = 36$). Ten studies reported on older adults selected for a specific disease or medical condition, such as Parkinson's disease ($k = 6$), post-surgical interventions for knees, hips, or backs ($k = 2$), postmenopausal women with osteopenia ($k = 1$), and chronic stroke ($k = 1$).

The duration and frequency of the exercise interventions ranged from 1 week to 1 year, 1–5 sessions/week, and 15–90 min/session. Of the 46 studies, 16 executed multicomponent (i.e., multifaceted) exercise interventions in at least one group. Reactive balance was assessed before and after the exercise interventions by use of laboratory-induced slip, trip, and falls, external impacts (e.g., pulling or pushing a body part), platform translation, and treadmill perturbation (e.g., rapid change of the speed) while participants were performing a steady-state task, such as standing or walking. Twenty studies provided training with a postural perturbation while standing or walking, and 11 of which implemented task-specific training (i.e., comparable reactive balance task included in the assessment and training) (Wolf et al., 1997; Bieryla et al., 2007; Beling and Roller, 2009; Mansfield et al., 2010; Parijat and Lockhart, 2012; Jagdhane et al., 2016; Morat et al., 2019; Okubo et al., 2019; Wang et al., 2019; Arghavani et al., 2020; Rieger et al., 2020). The characteristics of the participants and exercise interventions are presented in **Tables 2**, **3**, respectively. Outcome measures and main findings are summarized in **Supplementary Table 2**.

Risk of Bias

The summary of the risk of bias assessment across all included studies is presented in **Figure 2**. Detailed results of the assessment are reported in **Supplementary Table 3**. Overall, the majority of

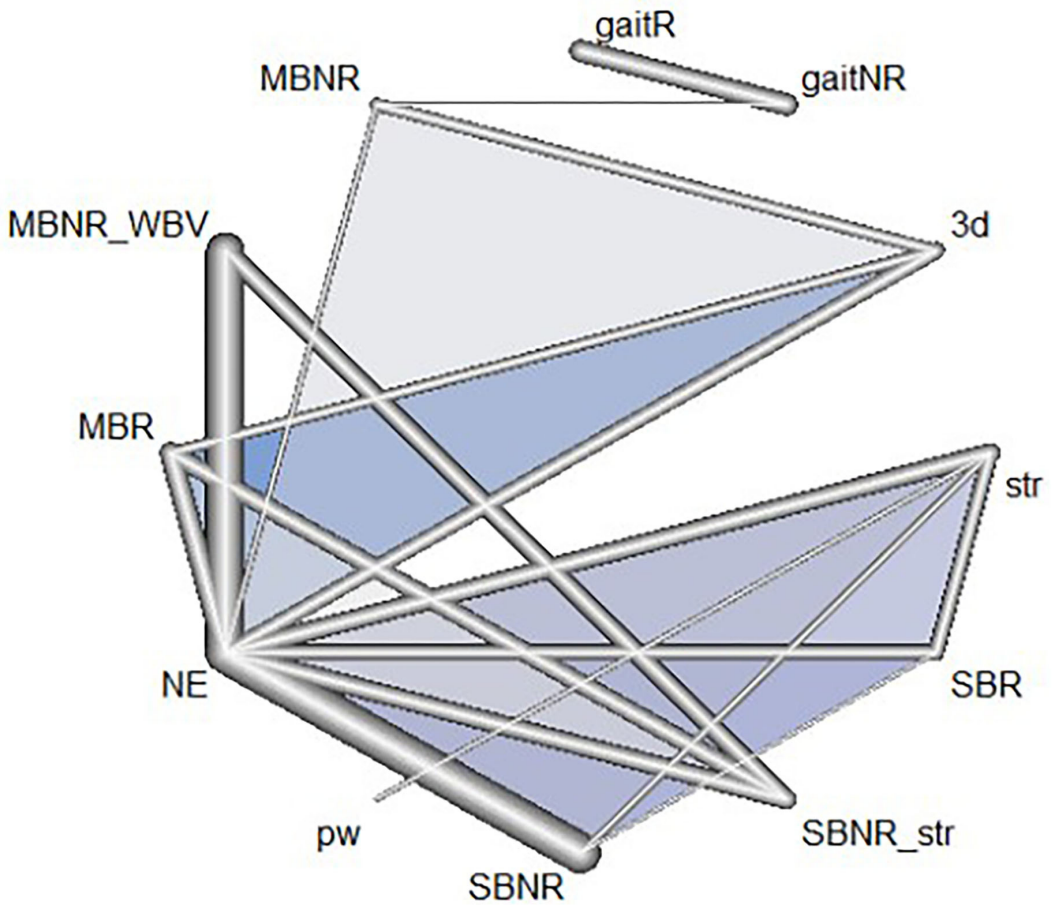


FIGURE 6 | Network geometry of the included exercise programs in healthy older adults: Each line indicates a direct comparison of two different exercise programs. The thickness of the edge is proportional to the number of direct comparisons in the network. Different exercise types combined in one program are connected via underscores. The blue triangles refer to multi-arm trials comprised of three exercise programs in the nodes. SBR, Single balance exercise including reactive balance component; SBNR, Single balance exercise not including reactive balance component; MBR, Multiple balance exercises including reactive balance component; MBNR, Multiple balance exercises not including reactive balance component; gaitR, Gait training including reactive balance component; gaitNR, Gait training not including reactive balance component; WBV, Whole body vibration; str, Strength; pw, Power; 3d, 3D exercise; NE, No exercise.

outcomes were at some concerns (50%) and high risk (48%), and only one study was rated as at low risk. Missing outcome data (46%) was the most influential source of high risk of bias. Selection of the reported result (83%), randomization process (76%), and deviations from intended interventions (61%) were also common sources of bias.

Network Meta-Analysis

Data from a total of 39 studies (n = 1388, age = 71.5 ± 3.9 years) were included in the NMA. Of the 15 exercise types reported in **Table 1**, 14 types were included in the NMA as functional training was implemented in only one study and consequently included in a disconnected network (Cabrera-Martos et al., 2020). There were 11 multi-arm trials, and three of which consisted of two groups sharing the same exercise type and the third group with another type (Hatzitaki et al., 2009; Ni et al., 2014; Lacroix et al., 2016); thus, data in these two groups were combined into a single group. Two exercise groups in studies by Gatts and Woollacott (2007) and Gatts (2008), str and NE groups in studies by Granacher et al. (2006, 2009), and two exercise groups in studies by Parijat et al. (2015a,b) shared the same participants, respectively. Thus, each of the aforementioned pairs of studies was combined as a single study in NMA. Overall, 17 exercise interventions with either single or multiple exercise components were included in the NMA. The geometric distribution of the network is depicted in **Figure 3**. When a study involves a trial arm with a combination of the pre-categorized exercise types, the combination was considered as another distinct exercise intervention.

Estimates of all exercise programs against all others in NMA were reported in a matrix (**Figure 4**). In the 17 exercise programs, SBR displayed the highest probability of being the most effective exercise intervention (SUCRA score = 0.90) for improving reactive balance, followed by pw (SUCRA score = 0.67) and gaitR (SUCRA score = 0.62) (**Table 4**).

TABLE 5 | Ranking of exercise interventions in healthy older adults.

Bayesian framework			Frequentist framework		
Ranking	**Exercise**	**SUCRA score**	**Ranking**	**Exercise**	**P-score**
1	SBR	0.90	1	SBR	0.95
2	pw	0.71	2	pw	0.76
3	str	0.52	3	str	0.53
4	gaitR	0.52	4	gaitR	0.52
5	SBNR	0.50	5	SBNR	0.52
6	MBR	0.47	6	MBR	0.47
7	MBNR	0.46	7	MBNR	0.46
8	MBNR + WBV	0.43	8	MBNR + WBV	0.41
9	SBNR + str	0.42	9	SBNR + str	0.41
10	gaitNR	0.40	10	gaitNR	0.37
11	3d	0.35	11	3d	0.32
12	NE	0.32	12	NE	0.28

SBR, Single balance exercise including reactive balance component; SBNR, Single balance exercise not including reactive balance component; MBR, Multiple balance exercises including reactive balance component; MBNR, Multiple balance exercises not including reactive balance component; gaitR, Gait training including reactive balance component; gaitNR, Gait training not including reactive balance component; WBV, Whole body vibration; str, Strength; pw, Power; 3d, 3D exercise; NE, No exercise.

The relative treatment effect estimates of each exercise program with the no-exercise program being the mutual contrast for comparison are presented in a forest plot (**Figure 5**). SBR, pw, and gaitR demonstrated the largest mean difference vs. NE; however SBR only demonstrated a statistically significant difference when compared to the no-exercise program (mean difference = 2.7, 95% CrI = 1.0 to 4.3). The trace plot, density plot, and Brooks-Gelman-Rubin diagnostic statistics showed good convergence, which signifies our data has converged to a reasonable distribution. Relatively reliable evidence was derived from the statistical consistency between direct and indirect evidence demonstrated by the node-splitting model ($p > 0.05$). According to the sensitivity analysis using a Frequentist framework of NMA, the ranking based on the P-scores showed identical results (**Table 4**). The results suggest that our main findings regarding the relative effectiveness of each exercise intervention are robust for future decisions.

Subgroup Analyses

In the subgroup analysis for healthy older adults ($k = 29$, $n = 1120$, age = 71.5 ± 3.7 years), effects of 12 exercise programs were compared (**Figure 6**).

According to the SUCRA scores, SBR was the highest-ranked exercise program (0.90), followed by pw (0.71), which was consistent with the ranking in the complete sample (**Table 5**). The other exercise programs ranked slightly differently from the NMA for the complete sample; however, the rankings based on the SUCRA scores were consistent with those estimated by P-scores in the frequentist framework (**Table 5**). The relative effects of all exercise interventions compared to NE were presented in **Figure 7**. A relative effect matrix was additionally created for all comparisons in the healthy older adults (**Figure 8**). Too few trials in other disease categories (Parkinson's disease: 6 trials, arthritis: 2 trials, osteopenia: 1 trial) and types of exercise interventions were available to establish a network in each category and conduct further disease-specific subgroup analyses.

For the second subgroup analysis regarding the types of reactive balance outcomes, the first type (simulated slip or trip while walking) was analyzed for gaitR vs. gaitNR using a multilevel MA due to insufficient trials in other treatment comparisons (i.e., only one direct comparison). The second (simulated forward falls), third (being pushed or pulled), and fourth (movable platform) types were analyzed using NMA. The fifth type (balance test battery) was not analyzed due to the insufficient number of exercise interventions and direct comparisons to establish a network. When a slip or trip was simulated while walking, participants showed greater improvements in measures of balance recoveries after gaitR training vs. gaitNR training (SMD = 0.60; 95% CI, 0.33 to 0.88). In other types, SBR presented the first or second highest probability of being the best intervention for improving each reactive balance task. The ranking and relative effects of each exercise vs. NE are reported in **Table 6** and **Figure 9**, respectively.

DISCUSSION

To our knowledge, this study is the first NMA to determine which type of exercise intervention is most effective in improving reactive balance in older adults. In this study, we compared the effects of commonly used exercise interventions on reactive balance in older adults. The NMA was used to analyze the data of 39 RCTs including 1,388 participants, which revealed that older adults receiving a balance exercise with a reactive balance component showed the most improvements in reactive balance, followed by power training (second) and gait training with a reactive balance component (third) among 17 different exercise interventions.

The results of this current study highlight the importance of applying the principle of specificity to training interventions designed to improve reactive balance. This is consistent with the hypothesis put forth by Grabiner et al. (2014) suggesting that task-specific perturbation training is superior to conventional exercise approaches in improving reactive balance capacity and thus preventing falls (Grabiner et al., 2014). Of the 46 trials in the current study, there were 20 trials including at least one exercise intervention with a reactive balance component, and ten of which utilized the same parameters of postural perturbations during the training and assessment (i.e., task-specific reactive balance training) (Wolf et al., 1997; Bieryla et al., 2007; Beling and Roller, 2009; Mansfield et al., 2010; Parijat and Lockhart, 2012; Morat et al., 2019; Okubo et al., 2019; Allin et al., 2020; Arghavani et al., 2020; Rieger et al., 2020). This latter point is especially important given that a specific type of reactive balance exercise has no, or at most a limited transfer effect on non-trained reactive balance tasks (Kümmel et al., 2016; Harper et al., 2021). The cognitive processes, muscle synergies, and succeeding kinematic strategies to counteract the perturbation

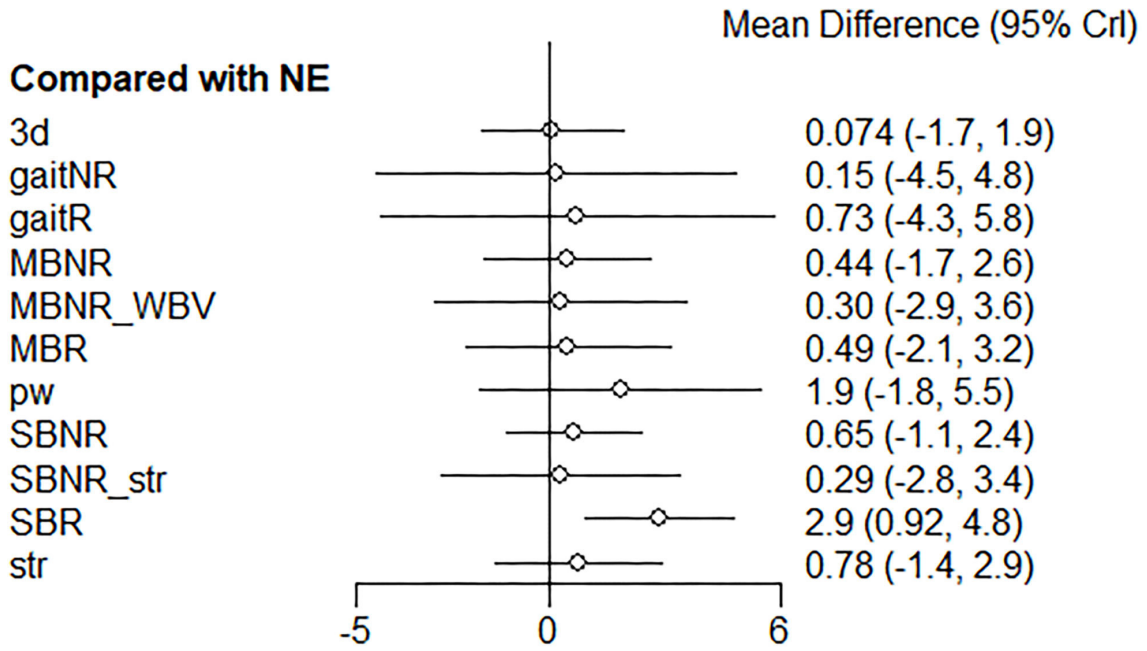


FIGURE 7 | Forest plot of the relative effects of exercise interventions with a no-exercise as a reference group in healthy older adults. SBR, Single balance exercise including reactive balance component; SBNR, Single balance exercise not including reactive balance component; MBR, Multiple balance exercises including reactive balance component; MBNR, Multiple balance exercises not including reactive balance component; gaitR, Gait training including reactive balance component; gaitNR, Gait training not including reactive balance component; WBV, Whole body vibration; str, Strength; pw, Power; 3d, 3D exercise; NE, No exercise.

are entirely determined by the parameters of the perturbations, such as type, magnitude, direction, and the point of application (Winter et al., 1990; Grabiner et al., 2014; Chen et al., 2017), and reactive balance improves in the tasks that are specifically trained with the same parameters. That is one of the reasons SBR showed greater improvements than other types of exercises. It seems reasonable to speculate that if all 20 trials used the exact same training and assessment tasks, the performance gains in reactive balance would be even greater. However, the estimates in the second subgroup analyses (i.e., types of reactive balance outcomes) regarding high SUCRA scores of SBR should be interpreted with caution given the small number of trials and several wide credible intervals.

VRepeated exposure to specific learning environments, therefore, leads to specific motor adaptation and learning. Motor adaptation is a learning process in which the nervous system learns how to predict and cancel impacts of a novel environment (e.g., perturbation), and ultimately maximize performance in that environment (Izawa et al., 2008). The cerebral cortex plays a key role in the acquisition and facilitation of balance recovery skills (Beck et al., 2007; Bolton, 2015). Through repeated exposure to a postural perturbation, our sensorimotor system learns (e.g., procedural learning) internal models for the sensorial prediction and motor commands and uses the learned models for an efficient and optimized movement plan (Izawa et al., 2008), ultimately improving compensatory reactions in older adults (Bohm et al., 2015; König et al., 2019). If mechanical perturbations transpire in consistent patterns with regards to the timing, magnitude, type, and direction, those who have undergone training using the same perturbation system could employ proactive (anticipatory) postural adjustments (i.e., feedforward control in anticipation of or before a postural perturbation) (Bhatt et al., 2006; Aruin et al., 2017; Curuk et al., 2020). Utilization of proactive postural adjustments, facilitated by the repeated exposure to the perturbation, significantly reduces the need for compensatory adjustments after a perturbation (Kanekar and Aruin, 2014); thus, the predictability regarding the perturbation and reactive balance task ultimately imparts greater adaptability and controllability. Such motor training is capable of altering corticospinal excitations and reorganizing motor maps and synaptic changes in the cerebral cortex, which ultimately facilitates the acquisition of a specific balance recovery skill (Beck et al., 2007; Grabiner et al., 2014), and the neuroplastic changes after training offer revealing clinical insights. However, when the patterns of a perturbation unpredictably change, the proactive postural adjustments, that are strictly relying on prior experience, can be deteriorated, which may compromise application to real-world falls where people rarely know in advance how and when they will get perturbed. Thus, a perturbation during training needs to be offered in various patterns to maximize the unpredictability and prepare older adults for the unpredictable nature of real-world falls (Harper et al., 2021). Further, to promote motor adaptation and learning, the elements of the training regimen should be properly determined first, and the challenge should be increased by adjusting the parameters of the perturbation, complexity of the context, and cognitive processing demands (Harper et al., 2021).

3d											
-0.058 (-4.65, 4.42)	gaitNR										
-0.63 (-5.60, 4.31)	-0.55 (-2.60, 1.45)	gaitR									
-0.37 (-2.23, 1.48)	-0.30 (-4.40, 3.85)	0.26 (-4.33, 4.85)	MBNR								
-0.26 (-3.99, 3.45)	-0.18 (-5.83, 5.53)	0.38 (-5.68, 6.43)	0.11 (-3.78, 4.02)	MBNR + WBV							
-0.44 (-3.34, 2.51)	-0.37 (-5.57, 4.91)	0.19 (-5.37, 5.83)	-0.06 (-3.29, 3.16)	-0.18 (-4.04, 3.71)	MBR						
0.07 (-1.75, 1.89)	0.15 (-4.53, 4.83)	0.70 (-4.40, 5.80)	0.44 (-1.74, 2.62)	0.33 (-2.94, 3.58)	0.51 (-2.16, 3.12)	NE					
-1.78 (-5.83, 2.25)	-1.72 (-7.59, 4.29)	-1.17 (-7.34, 5.19)	-1.41 (-5.65, 2.83)	-1.54 (-6.34, 3.38)	-1.35 (-5.83, 3.12)	-1.86 (-5.45, 1.76)	pw				
-0.57 (-3.08, 1.94)	-0.49 (-5.47, 4.53)	0.06 (-5.27, 5.46)	-0.20 (-2.96, 2.55)	-0.31 (-4.02, 3.36)	-0.14 (-3.33, 3.06)	-0.65 (-2.39, 1.11)	1.20 (-2.64, 5.05)	SBNR			
-0.25 (-3.75, 3.31)	-0.15 (-5.72, 5.43)	0.41 (-5.49, 6.34)	0.13 (-3.59, 3.92)	0.03 (-3.21, 3.28)	0.20 (-3.22, 3.59)	-0.30 (-3.40, 2.82)	1.56 (-3.21, 6.35)	0.34 (-3.20, 3.95)	SBNR + str		
-2.80 (-5.44, -0.14)	-2.72 (-7.74, 2.38)	-2.16 (-7.59, 3.35)	-2.43 (-5.33, 0.48)	-2.54 (-6.31, 1.26)	-2.36 (-5.65, 0.92)	-2.87 (-4.81, -0.94)	-1.00 (-4.92, 2.90)	-2.23 (-4.59, 0.18)	-2.57 (-6.21, 1.10)	SBR	
-0.70 (-3.55, 2.11)	-0.62 (-5.79, 4.59)	-0.06 (-5.63, 5.52)	-0.33 (-3.41, 2.74)	-0.44 (-4.38, 3.46)	-0.27 (-3.69, 3.15)	-0.78 (-2.94, 1.40)	1.08 (-1.80, 3.97)	-0.13 (-2.67, 2.40)	-0.47 (-4.27, 3.34)	2.10 (-0.55, 4.71)	str

FIGURE 8 | Relative effect estimates with 95% credible intervals of all pairs of exercise interventions in healthy older adults.

The greatest effect of SBR and relatively less effective multicomponent exercise interventions can be further scrutinized via several critical principles of exercise training including volume, intensity, and frequency. Training volume is largely determined by the time commitment (duration) of the training. However, the total duration and frequency of the interventions are broadly ranged across the included studies as previously reported, and the average duration of each training session was 52.2 ± 19.7 min. If an intervention included multiple types of exercises in a single session, the intervention may lack the critical time needed to focus on reactive balance training. According to Burgomaster et al. (Burgomaster et al., 2008), low-volume, high-intensity training and high-volume, low-intensity training induce comparable changes in selected whole-body and skeletal muscle adaptations when the frequencies and the total durations are identical (Burgomaster et al., 2008; Hawley, 2008). Thus, if lack of time is a barrier to satisfying the need for reactive balance training, the intensity aspect of the training should be considered as a way to compensate for the deficit and induce targeted changes in reactive balance. Further, it is encouraging that Bhatt and Pai have demonstrated significant improvements in reactive balance performance after a single high-intensity training with task-specific postural perturbations (Bhatt and Pai, 2009). This is particularly noteworthy given that such minimal training effects were retained for several months when properly selecting the intensity and specificity of the training despite the relatively small total volume. Thus, future trials may wish to take account of the aforementioned factors, including specificity, volume, and intensity of the training to maximize the time-effective transfer to real-world scenarios.

Lastly, given the high ranking of power training, the probable inter-relation with reactive balance is clinically notable. In situations where a mechanical perturbation is applied and a fall begins, the rate of torque development in the lower or upper extremity joints with intersegment coordination has

TABLE 6 | Ranking of exercise interventions in each reactive balance outcome category.

A			B			C		
Ranking	**Exercise**	**SUCRA score**	**Ranking**	**Exercise**	**SUCRA score**	**Ranking**	**Exercise**	**SUCRA score**
1	MBNR + WBV	0.77	1	SBR	0.73	1	SBR	0.79
2	SBR	0.65	2	SBNR + flex	0.63	2	MBR	0.75
3	Str	0.64	3	3d	0.35	3	pw	0.72
4	SBNR + str	0.52	4	NE	0.30	4	balUS	0.60
5	pw	0.39				5	str	0.58
6	MBR	0.39				6	MBR +gaitNR	0.54
7	NE	0.14				7	MBNR	0.48
						8	MBNR +gaitNR	0.48
						9	SBNR	0.43
						10	SBNR + flex	0.43
						11	3d	0.30
						12	MBNR + WBV	0.25
						13	NE	0.14

A. Simulated forward falls, B. Being pushed or pulled, C. Movable platform. SBR, Single balance exercise including reactive balance component; SBNR, Single balance exercise not including reactive balance component; MBR, Multiple balance exercises including reactive balance component; MBNR, Multiple balance exercises not including reactive balance component; WBV, Whole body vibration; str, Strength; pw, Power; 3d, 3D exercise; NE, No exercise.

been considered as a critical determinant of balance recovery by taking a step or reach to grasp (Madigan, 2006). Aging inherently brings a loss of motor neurons, associated with apoptosis, and reduction and denervation of muscle fibers, specifically related to type II muscle fibers. These changes lead to a decrease in the muscles' capacity to produce maximum muscle strength, power, and rate of force development (Aagaard et al., 2010). In general, fallers generate less muscle power than non-fallers, and older adults generate less power than young adults (Madigan, 2006; Perry et al., 2007). By utilizing the comparability between muscle power and reactive balance, such as forceful and controlled movements with high velocity, all power training groups in the current analysis demonstrated improvements in measures of reactive balance. There are a handful of studies investigating the correlations between muscle power and reactive balance performances (Muehlbauer et al., 2015); however, the effectiveness of power training on reactive balance has been explored only in a few, recent trials (Pamukoff et al., 2014; Inacio et al., 2018; Cherup et al., 2019). The results of this study may have implications for future directions in assessing the relationship between muscle power and reactive balance.

Given that the vast majority of falls occur while walking (Tinetti et al., 1988; Berg et al., 1997; Li et al., 2006; Kelsey et al., 2012; Robinovitch et al., 2013), training to counter postural perturbations while walking is imperative. However, the ranking of gaitR was relatively lower than SBR. Because gaitR was only compared with gaitNR in the network, the ranking was dominantly determined by indirect evidence. Standard errors from indirect evidence are greater than those from direct evidence, which represents the lower accuracy of an estimate (Higgins et al., 2019). Thus, the indirect evidence should be interpreted with caution, and more RCTs with direct comparisons between gaitR and other exercises may guarantee more accurate posterior distributions and the ranking of gaitR.

Clinical Implications

Considering the findings of this study, it would be advisable for clinicians to preferentially include reactive balance training in line with specifically targeted context, direction, and type of postural perturbations, and power training as a secondarily or complementary approach to improve reactive balance in older adults irrespective of their clinical classifications. Multicomponent exercise interventions not including a reactive balance component may not bring as marked changes in reactive balance as a single reactive balance training does, whereas they still have benefits regarding general health and physical functioning. The possibility of task-specific training adaptations with balance training using external mechanical perturbations has far-reaching clinical and research implications. In fact, beyond simply training one specific type of balance reaction (e.g., a slip), future trials may wish to include multiple types of reactive

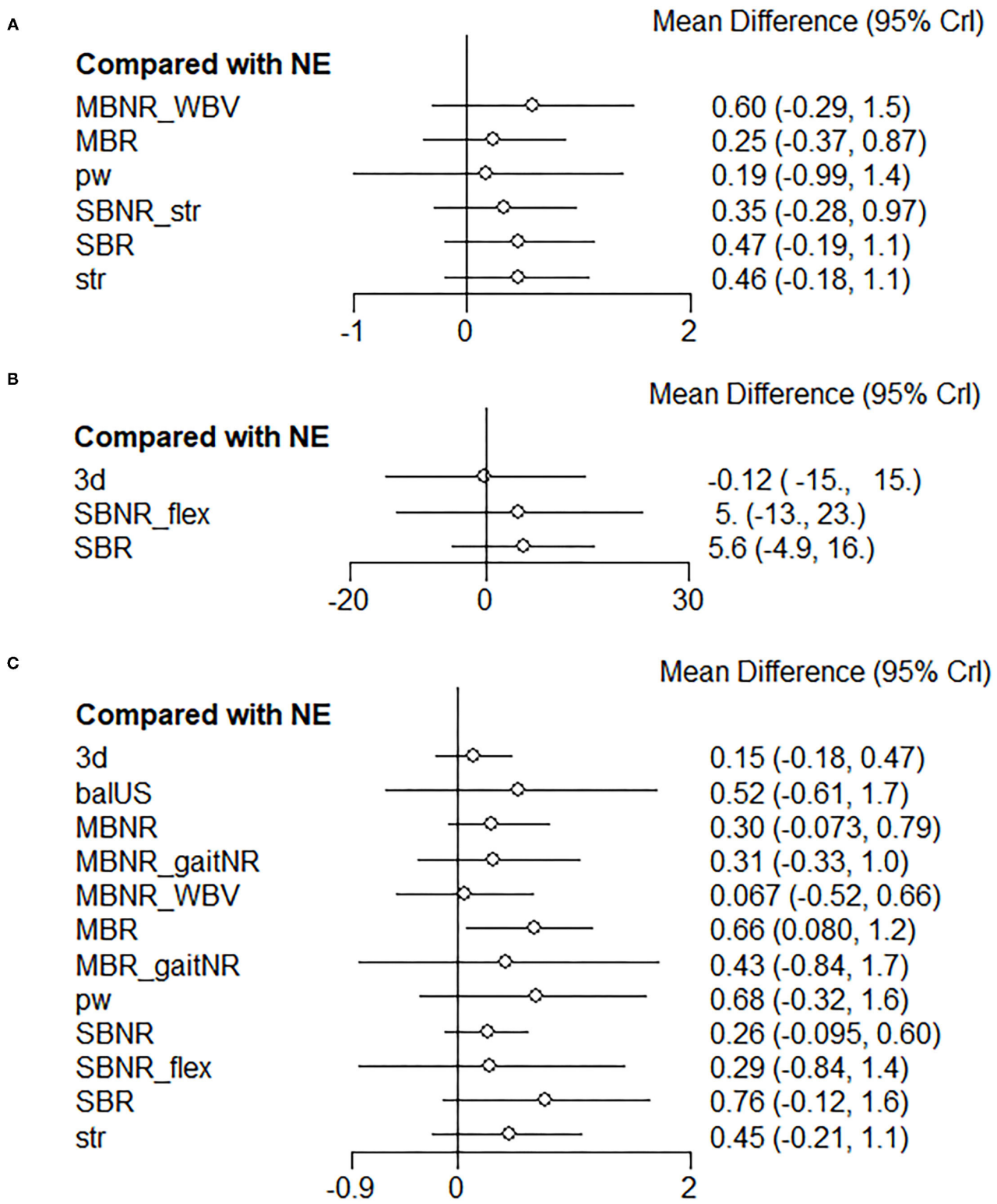


FIGURE 9 | Forest plots of the relative effects of exercise interventions with a no-exercise as a reference group in each reactive balance outcome category. **(A)** Simulated forward falls, **(B)** Being pushed or pulled, and **(C)** Movable platform. SBR, Single balance exercise including reactive balance component; SBNR, Single balance exercise not including reactive balance component; MBR, Multiple balance exercises including reactive balance component; MBNR, Multiple balance exercises not including reactive balance component; WBV, Whole body vibration; str, Strength; pw, Power; 3d, 3D exercise; NE, No exercise.

balance tasks in various simulated contexts that are likely to occur in daily life and appraise the generalizability and ecological validity of the trained tasks from a long-term perspective. Moreover, the addition of power training may synergize the effects on functional reflex activities as well as general functional capabilities needed for daily tasks and reducing falls in older adults.

Strengths and Limitations

One of the major strengths of the current study is the use of a NMA. The notable advantage of a NMA over a conventional pairwise meta-analysis is the ability to allow for indirect comparisons, accounting for the effects of multiple interventions in a single statistical model (Schwarzer et al., 2015). Thus, a NMA concurrently summarizes both direct and indirect comparisons between multifarious interventions and enables more complex statistical models and broader interpretation. Random-effects models attempt to generalize the results beyond the trials included in the NMA with an assumption that the selected trials are random samples from a larger population (Cheung et al., 2012). Accordingly, the use of a NMA with a random-effects model in this study enhances the applicability and generalizability of study findings. It should be noted that in general, the indirect estimates tend to have greater variance than direct estimates, and the reliability of the indirect estimates are influenced by the number of direct estimates in the network (Dias et al., 2018). Future meta-analyses may wish to assess publication bias and heterogeneity with a greater number of trials in each direct comparison.

The interpretations of the results in the current study are limited due to small sample sizes and the existence of the probable risk of bias in the included studies. For example, only two trials included more than 100 total participants (Bogaerts et al., 2007; Wang et al., 2019). Furthermore, there was heterogeneity in participants and exercise interventions. For example, there were several distinct disease groups, and the frequency and duration were set differently for various exercise interventions pooled together. With further trials, future reviews may wish to break down the analyses on the basis of hypothetical effect modifiers, such as detailed age and disease groups, baseline functional capacities, or dosage of intervention, for more specific clinical decisions. Also, the low number of trials per comparison precluded investigating sources of publication bias and heterogeneity, and the overall risk of bias was appraised as some concern or high-risk level. Thus, a comprehensive search of published and unpublished works of literature with a paired screening process was conducted to guarantee all available literature was identified to reduce the potential risk of publication bias. Considering the number of trials per each direct comparison, sample sizes, and overall risks of bias, thei results of our analyses may as such guide future research. Despite the aforementioned limitations, we believe that this systematic review with a NMA shed light on better understanding effective interventions for reactive balance in older adults via more comprehensive and inclusive analyses of available literature.

CONCLUSIONS

In conclusion, our NMA indicates that SBR, which simulates a real-life fall scenario and induces a specific balance recovery, is generally more efficacious in improving reactive balance than any other exercise intervention in older adults. Importantly, power training also appears to have greater impacts on reactive balance than other exercise interventions. Our results highlight the importance of task-specific exercise interventions with respect to the targeted postural perturbation and reactions. More trials with high methodological quality, low risk of bias, larger samples, and older adults with a specific disease or disability need to be conducted to construct a comprehensive literature basis, which would facilitate a more thorough NMA. The findings of this study could be used to design exercise-based interventions for improving reactive balance in older adults.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

YK, MV, and EB performed the search, selection, data extraction, and risk of bias assessments. YK performed all data analyses, drafted the manuscript, and attests that all listed authors meet the authorship criteria and that no others meeting the criteria have been omitted. All authors participated in the conception and design of the study, provided feedback on early and advanced drafts of the manuscript, critically revised for important intellectual content, approved the final version of the manuscript, and agreed to be accountable for all aspects of the work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2021.764826/full#supplementary-material

REFERENCES

Aagaard, P., Suetta, C., Caserotti, P., Magnusson, S. P., and Kjær, M. (2010). Role of the nervous system in sarcopenia and muscle atrophy with aging: strength training as a countermeasure. *Scandinav. J. Med. Sci. Sport.* 20, 49–64. doi: 10.1111/j.1600-0838.2009.01084.x

Alissa, N., Akinlosotu, R. Y., Shipper, A. G., Wheeler, L. A., and Westlake, K. P. (2020). A systematic review of upper extremity responses during

reactive balance perturbations in aging. *Gait. Posture.* 82, 138–146. doi: 10.1016/j.gaitpost.2020.08.134

Allin, L. J., Brolinson, P. G., Beach, B. M., Kim, S., Nussbaum, M. A., Roberto, K. A., et al. (2020). Perturbation-based balance training targeting both slip- and trip-induced falls among older adults: a randomized controlled trial. *BMC Geriatr.* 20, 205. doi: 10.1186/s12877-020-01605-9

Ambrose, A. F., Paul, G., and Hausdorff, J. M. (2013). Risk factors for falls among older adults: A review of the literature. *Maturitas.* 75, 51–61. doi: 10.1016/j.maturitas.2013.02.009

Arampatzis, A., Peper, A., and Bierbaum, S. (2011). Exercise of mechanisms for dynamic stability control increases stability performance in the elderly. *J. Biomech.* 44, 52–58. doi: 10.1016/j.jbiomech.2010.08.023

Arghavani, H., Zolaktaf, V., and Lenjannejadian, S. (2020). Comparing the effects of anticipatory postural adjustments focused training and balance training on postural preparation, balance confidence and quality of life in elderly with history of a fall. *Aging Clin. Exper. Res.* 32, 1757–1765. doi: 10.1007/s40520-019-01358-5

Aruin, A. S., Ganesan, M., and Lee, Y. (2017). Improvement of postural control in individuals with multiple sclerosis after a single-session of ball throwing exercise. *Mult. Scleros. Relat. Disord.* 17, 224–229. doi: 10.1016/j.msard.2017.08.013

Beck, S., Taube, W., Gruber, M., Amtage, F., Gollhofer, A., and Schubert, M. (2007). Task-specific changes in motor evoked potentials of lower limb muscles after different training interventions. *Brain Res.* 1179, 51–60. doi: 10.1016/j.brainres.2007.08.048

Beling, J., and Roller, M. (2009). Multifactorial intervention with balance training as a core component among fall-prone older adults. *J. Geriatr. Phys. Ther.* 32, 125–133. doi: 10.1519/00139143-200932030-00008

Berg, W. P., Alessio, H. M., Mills, E. M., and Tong, C. (1997). Circumstances and consequences of falls in independent community-dwelling older adults. *Age Ageing.* 26, 261–268. doi: 10.1093/ageing/26.4.261

Bhatt, T., and Pai, Y.-C. (2009). Prevention of slip-related backward balance loss: the effect of session intensity and frequency on long-term retention. *Arch. Phys. Med. Rehabil.* 90, 34–42. doi: 10.1016/j.apmr.2008.06.021

Bhatt, T., Wening, J. D., and Pai, Y.-C. (2006). Adaptive control of gait stability in reducing slip-related backward loss of balance. *Exp. Brain Res.* 170, 61–73. doi: 10.1007/s00221-005-0189-5

Bieryla, K. A., Madigan, M. L., and Nussbaum, M. A. (2007). Practicing recovery from a simulated trip improves recovery kinematics after an actual trip. *Gait. Posture.* 26, 208–213. doi: 10.1016/j.gaitpost.2006.09.010

Bogaerts, A., Verschueren, S., Delecluse, C., Claessens, A. L., and Boonen, S. (2007). Effects of whole body vibration training on postural control in older individuals: a 1 year randomized controlled trial. *Gait. Posture.* 26, 309–316. doi: 10.1016/j.gaitpost.2006.09.078

Bohm, S., Mademli, L., Mersmann, F., and Arampatzis, A. (2015). Predictive and reactive locomotor adaptability in healthy elderly: A systematic review and meta-analysis. *Sports Med.* 45, 1759–1777. doi: 10.1007/s40279-015-0413-9

Bolton, D. A. E. (2015). The role of the cerebral cortex in postural responses to externally induced perturbations. *Neurosci. Biobehav Rev.* 57, 142–155. doi: 10.1016/j.neubiorev.2015.08.014

Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2009). *Introduction to Meta-Analysis*. John Wiley and Sons. doi: 10.1002/9780470743386

Brooks, S. P., and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Null.* 7, 434–455. doi: 10.1080/10618600.1998.10474787

Bruderer-Hofstetter, M., Rausch-Osthoff, A.-K., Meichtry, A., Münzer, T., and Niedermann, K. (2018). Effective multicomponent interventions in comparison to active control and no interventions on physical capacity, cognitive function and instrumental activities of daily living in elderly people with and without mild impaired cognition – A systematic review and network meta-analysis. *Age. Res. Rev.* 45, 1–14. doi: 10.1016/j.arr.2018.04.002

Burgomaster, K. A., Howarth, K. R., Phillips, S. M., Rakobowchuk, M., MacDonald, M. J., McGee, S. L., et al. (2008). Similar metabolic adaptations during exercise after low volume sprint interval and traditional endurance training in humans. *J. Physiol.* 586, 151–160. doi: 10.1113/jphysiol.2007.142109

Cabrera-Martos, I., Jiménez-Martín, A. T., López-López, L., Rodríguez-Torres, J., Ortiz-Rubio, A., and Valenza, M. C. (2020). Effects of a core stabilization training program on balance ability in persons with Parkinson's disease: a randomized controlled trial. *Clin. Rehabil.* 34, 764–772. doi: 10.1177/0269215520918631

Chen, B., Lee, Y.-J., and Aruin, A. S. (2017). Role of point of application of perturbation in control of vertical posture. *Exp. Brain Res.* 235, 3449–3457. doi: 10.1007/s00221-017-5069-2

Cherup, N. P., Buskard, A. N. L., Strand, K. L., Roberson, K. B., Michiels, E. R., Kuhn, J. E., et al. (2019). Power vs strength training to improve muscular strength, power, balance and functional movement in individuals diagnosed with Parkinson's disease. *Exp. Gerontol.* 128, 110740. doi: 10.1016/j.exger.2019.110740

Cheung, M. W.-L., Ho, R. C. M., Lim, Y., and Mak, A. (2012). Conducting a meta-analysis: basics and good practices. *Int. J. Rheumat. Dis.* 15, 129–135. doi: 10.1111/j.1756-185X.2012.01712.x

Chyu, M.-C., James, C. R., Sawyer, S. F., Brismée, J.-M., Xu, K. T., Poklikuha, G., et al. (2010). Effects of tai chi exercise on posturography, gait, physical function and quality of life in postmenopausal women with osteopaenia: A randomized clinical study. *Clin. Rehabil.* 24, 1080–1090. doi: 10.1177/0269215510375902

Curuk, E., Lee, Y., and Aruin, A. S. (2020). Individuals with stroke improve anticipatory postural adjustments after a single session of targeted exercises. *Human Movem. Sci.* 69, 102559. doi: 10.1016/j.humov.2019.102559

Deandrea, S., Lucenteforte, E., Bravi, F., Foschi, R., La Vecchia, C., and Negri, E. (2010). Risk factors for falls in community-dwelling older people: a systematic review and meta-analysis. *Epidemiology.* 21, 658–668. doi: 10.1097/EDE.0b013e3181e89905

Dias, S., Ades, A. E., Welton, N. J., Jansen, J. P., and Sutton, A. J. (2018). *Network Meta-Analysis for Decision-Making*. John Wiley and Sons. doi: 10.1002/9781118951651

Donath, L., Roth, R., Hürlimann, C., Zahner, L., and Faude, O. (2016). Pilates vs. balance training in healthy community-dwelling seniors: A 3-arm, randomized controlled trial. *Int. J. Sports Med.* 37, 202–210. doi: 10.1055/s-0035-1559695

Fu, R., Vandermeer, B. W., Shamliyan, T. A., O'Neil, M. E., Yazdi, F., Fox, S. H., et al. (2008). Handling continuous outcomes in quantitative synthesis. In *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* AHRQ Methods for Effective Health Care. (Rockville (MD): Agency for Healthcare Research and Quality (US)). Available online at: http://www.ncbi.nlm.nih.gov/books/NBK154408/. (Accessed May 25, 2021).

Gatts, S. (2008). Neural mechanisms underlying balance control in Tai Chi. *Med. Sport Sci.* 52, 87–103. doi: 10.1159/000134289

Gatts, S. K., and Woollacott, M. H. (2007). How Tai Chi improves balance: biomechanics of recovery to a walking slip in impaired seniors. *Gait. Posture.* 25, 205–214. doi: 10.1016/j.gaitpost.2006.03.011

Grabiner, M. D., Crenshaw, J. R., Hurt, C. P., Rosenblatt, N. J., and Troy, K. L. (2014). Exercise-Based Fall Prevention: Can You Be a Bit More Specific? *Exer. Sport Sci. Rev.* 42, 161–168. doi: 10.1249/JES.0000000000000023

Granacher, U., Gollhofer, A., and Strass, D. (2006). Training induced adaptations in characteristics of postural reflexes in elderly men. *Gait. Posture.* 24, 459–466. doi: 10.1016/j.gaitpost.2005.12.007

Granacher, U., Gruber, M., and Gollhofer, A. (2009). Resistance training and neuromuscular performance in seniors. *Int. J. Sports Med.* 30, 652–657. doi: 10.1055/s-0029-1224178

Hamed, A., Bohm, S., Mersmann, F., and Arampatzis, A. (2018). Exercises of dynamic stability under unstable conditions increase muscle strength and balance ability in the elderly. *Scand. J. Med. Sci. Sports* 28, 961–971. doi: 10.1111/sms.13019

Harper, S. A., Beethe, A. Z., Dakin, C. J., and Bolton, D. A. E. (2021). Promoting Generalized Learning in Balance Recovery Interventions. *Brain Sci.* 11, 402. doi: 10.3390/brainsci11030402

Hatzitaki, V., Voudouris, D., Nikodelis, T., and Amiridis, I. G. (2009). Visual feedback training improves postural adjustments associated with moving obstacle avoidance in elderly women. *Gait. Posture.* 29, 296–299. doi: 10.1016/j.gaitpost.2008.09.011

Hawley, J. A. (2008). Specificity of training adaptation: time for a rethink? *J. Physiol.* 586, 1–2. doi: 10.1113/jphysiol.2007.147397

Hespanhol, L., Vallio, C. S., Costa, L. M., and Saragiotto, B. T. (2019). Understanding and interpreting confidence and credible intervals around effect estimates. *Braz. J. Phys. Ther.* 23, 290–301. doi: 10.1016/j.bjpt.2018.12.006

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., et al. (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley and Sons. doi: 10.1002/9781119536604

Howe, T. E., Rochester, L., Neil, F., Skelton, D. A., and Ballinger, C. (2011). Exercise for improving balance in older people. *Cochrane Datab. System. Rev.* doi: 10.1002/14651858.CD004963.pub3

Hu, M. H., and Woollacott, M. H. (1994). Multisensory training of standing balance in older adults: II. Kinematic and electromyographic postural responses. *J. Gerontol.* 49, M62–71. doi: 10.1093/geronj/49.2.m62

Hutton, B., Salanti, G., Caldwell, D. M., Chaimani, A., Schmid, C. H., Cameron, C., et al. (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann. Intern. Med.* 162, 777–784. doi: 10.7326/M14-2385

Inacio, M., Creath, R., and Rogers, M. W. (2018). Low-dose hip abductor-adductor power training improves neuromechanical weight-transfer control during lateral balance recovery in older adults. *Clin. Biomech.* 60, 127–133. doi: 10.1016/j.clinbiomech.2018.10.018

Izawa, J., Rane, T., Donchin, O., and Shadmehr, R. (2008). Motor adaptation as a process of reoptimization. *J. Neurosci.* 28, 2883–2891. doi: 10.1523/JNEUROSCI.5359-07.2008

Jagdhane, S., Kanekar, N., and, S., Aruin, A. (2016). The effect of a four-week balance training program on anticipatory postural adjustments in older adults: A pilot feasibility study. *Curr. Aging Sci.* 9, 295–300. doi: 10.2174/1874609809666160413113443

Kanekar, N., and Aruin, A. S. (2014). Aging and balance control in response to external perturbations: role of anticipatory and compensatory postural mechanisms. *Age (Dordr)* 36, 9621. doi: 10.1007/s11357-014-9621-8

Kelsey, J. L., Procter-Gray, E., Hannan, M. T., and Li, W. (2012). Heterogeneity of falls among older adults: implications for public health prevention. *Am. J. Public Health.* 102, 2149–2156. doi: 10.2105/AJPH.2012.300677

Kim, S., and Lockhart, T. (2010). Effects of 8 weeks of balance or weight training for the independently living elderly on the outcomes of induced slips. *Int. J. Rehabil. Res.* 33, 49–55. doi: 10.1097/MRR.0b013e32832e6b5e

Kim, Y., Vakula, M. N., Waller, B., and Bressel, E. (2020). A systematic review and meta-analysis comparing the effect of aquatic and land exercise on dynamic balance in older adults. *BMC Geriatr.* 20, 302. doi: 10.1186/s12877-020-01702-9

Klamroth, S., Gaßner, H., Winkler, J., Eskofier, B., Klucken, J., Pfeifer, K., et al. (2019). Interindividual balance adaptations in response to perturbation treadmill training in persons with Parkinson disease. *J. Neurol. Phys. Ther.* 43, 224–232. doi: 10.1097/NPT.0000000000000291

König, M., Epro, G., Seeley, J., Potthast, W., and Karamanidis, K. (2019). Retention and generalizability of balance recovery response adaptations from trip perturbations across the adult life span. *J. Neurophysiol.* 122, 1884–1893. doi: 10.1152/jn.00380.2019

Kümmel, J., Kramer, A., Giboin, L.-S., and Gruber, M. (2016). Specificity of balance training in healthy individuals: a systematic review and meta-analysis. *Sports Med.* 46, 1261–1271. doi: 10.1007/s40279-016-0515-z

Lacroix, A., Kressig, R. W., Muehlbauer, T., Gschwind, Y. J., Pfenninger, B., Bruegger, O., et al. (2016). Effects of a supervised versus an unsupervised combined balance and strength training program on balance and muscle power in healthy older adults: a randomized controlled trial. *GER.* 62, 275–288. doi: 10.1159/000442087

Lai, C.-C., Tu, Y.-K., Wang, T.-G., Huang, Y.-T., and Chien, K.-L. (2018). Effects of resistance training, endurance training and whole-body vibration on lean body mass, muscle strength and physical performance in older people: a systematic review and network meta-analysis. *Age Age.* 47, 367–373. doi: 10.1093/ageing/afy009

Lesinski, M., Hortobágyi, T., Muehlbauer, T., Gollhofer, A., and Granacher, U. (2015). Effects of balance training on balance performance in healthy older adults: a systematic review and meta-analysis. *Sports Med.* 45, 1721–1738. doi: 10.1007/s40279-015-0375-y

Li, J. X., Xu, D. Q., and Hong, Y. (2009). Changes in muscle strength, endurance, and reaction of the lower extremities with Tai Chi intervention. *J. Biomech.* 42, 967–971. doi: 10.1016/j.jbiomech.2009.03.001

Li, W., Keegan, T. H. M., Sternfeld, B., Sidney, S., Quesenberry, C. P., and Kelsey, J. L. (2006). Outdoor falls among middle-aged and older adults: a neglected public health problem. *Am. J. Public Health.* 96, 1192–1200. doi: 10.2105/AJPH.2005.083055

Lunn, D. J., Thomas, A., Best, N., and Spiegelhalter, D. (2000). WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statist. Comput.* 10, 325–337. doi: 10.1023/A:1008929526011

Ma, A. W. W., Wang, H.-K., Chen, D.-R., Chen, Y.-M., Chak, Y. T. C., Chan, J. W. Y., et al. (2019). Chinese martial art training failed to improve balance or inhibit falls in older adults. *Percept. Mot. Skills* 126, 389–409. doi: 10.1177/0031512518824945

Madigan, M. L. (2006). Age-related differences in muscle power during single-step balance recovery. *J. Appl. Biomech.* 22, 186–193. doi: 10.1123/jab.22.3.186

Mahoney, J. R., Cotton, K., and Verghese, J. (2019). Multisensory integration predicts balance and falls in older adults. *J. Gerontol. Series A* 74, 1429–1435. doi: 10.1093/gerona/gly245

Mansfield, A., Peters, A. L., Liu, B. A., and Maki, B. E. (2010). Effect of a perturbation-based balance training program on compensatory stepping and grasping reactions in older adults: a randomized controlled trial. *Phys. Ther.* 90, 476–491. doi: 10.2522/ptj.20090070

Marigold, D. S., Eng, J. J., Dawson, A. S., Inglis, J. T., Harris, J. E., and Gylfadóttir, S. (2005). Exercise leads to faster postural reflexes, improved balance and mobility, and fewer falls in older persons with chronic stroke. *J. Am. Geriatr. Soc.* 53, 416–423. doi: 10.1111/j.1532-5415.2005.53158.x

McCrum, C., Gerards, M. H. G., Karamanidis, K., Zijlstra, W., and Meijer, K. (2017). A systematic review of gait perturbation paradigms for improving reactive stepping responses and falls risk among healthy older adults. *Eur. Rev. Ag. Phys. Activity.* 14, 3. doi: 10.1186/s11556-017-0173-7

Moore, B. M., Adams, J. T., Willcox, S., and Nicholson, J. (2019). The effect of active physical training interventions on reactive postural responses in older adults: a systematic review. *J. Aging Phys. Activity.* 27, 252–264. doi: 10.1123/japa.2017-0347

Morat, M., Bakker, J., Hammes, V., Morat, T., Giannouli, E., Zijlstra, W., et al. (2019). Effects of stepping exergames under stable versus unstable conditions on balance and strength in healthy community-dwelling older adults: A three-armed randomized controlled trial. *Exper. Gerontol.* 127, 110719. doi: 10.1016/j.exger.2019.110719

Muehlbauer, T., Gollhofer, A., and Granacher, U. (2015). Associations between measures of balance and lower-extremity muscle strength/power in healthy individuals across the lifespan: a systematic review and meta-analysis. *Sports Med.* 45, 1671–1692. doi: 10.1007/s40279-015-0390-z

Ni, M., Mooney, K., Richards, L., Balachandran, A., Sun, M., Harriell, K., et al. (2014). Comparative impacts of Tai Chi, balance training, and a specially-designed yoga program on balance in older fallers. *Arch. Phys. Med. Rehabil.* 95, 1620–1628.e30. doi: 10.1016/j.apmr.2014.04.022

Ochi, A., Abe, T., Yamada, K., Ibuki, S., Tateuchi, H., and Ichihashi, N. (2015). Effect of balance exercise in combination with whole-body vibration on muscle activity of the stepping limb during a forward fall in older women: a randomized controlled pilot study. *Arch. Gerontol. Geriatr.* 60, 244–251. doi: 10.1016/j.archger.2014.11.011

Okubo, Y., Schoene, D., Caetano, M. J., Pliner, E. M., Osuka, Y., Toson, B., et al. (2021). Stepping impairment and falls in older adults: a systematic review and meta-analysis of volitional and reactive step tests. *Ageing Res. Rev.* 66, 101238. doi: 10.1016/j.arr.2020.101238

Okubo, Y., Sturnieks, D. L., Brodie, M. A., Duran, L., and Lord, S. R. (2019). Effect of reactive balance training involving repeated slips and trips on balance recovery among older adults: a blinded randomized controlled trial. *J. Gerontol.* 74, 1489–1496. doi: 10.1093/gerona/glz021

Osoba, M. Y., Rao, A. K., Agrawal, S. K., and Lalwani, A. K. (2019). Balance and gait in the elderly: a contemporary review. *Laryngosc. Invest. Otolaryngol.* 4, 143–153. doi: 10.1002/lio2.252

Pamukoff, D. N., Haakonssen, E. C., Zaccaria, J. A., Madigan, M. L., Miller, M. E., and Marsh, A. P. (2014). The effects of strength and power training on single-step balance recovery in older adults: a preliminary study. *Clin. Interv. Aging.* 9, 697–704. doi: 10.2147/CIA.S59310

Parijat, P., and Lockhart, T. E. (2012). Effects of moveable platform training in preventing slip-induced falls in older adults. *Ann. Biomed. Eng.* 40, 1111–1121. doi: 10.1007/s10439-011-0477-0

Parijat, P., Lockhart, T. E., and Liu, J. (2015a). Effects of perturbation-based slip training using a virtual reality environment on slip-induced falls. *Ann. Biomed. Eng.* 43, 958–967. doi: 10.1007/s10439-014-1128-z

Parijat, P., Lockhart, T. E., and Liu, J. (2015b). EMG and kinematic responses to unexpected slips after slip training in virtual reality. *IEEE Transac. On Bio-Med. Eng.* 62, 593–599. doi: 10.1109/TBME.2014.2361324

Perry, M. C., Carville, S. F., Smith, I. C. H., Rutherford, O. M., and Newham, D. J. (2007). Strength, power output and symmetry of leg muscles: effect of age and history of falling. *Eur. J. Appl. Physiol.* 100, 553–561. doi: 10.1007/s00421-006-0247-0

Pluchino, A., Lee, S. Y., Asfour, S., Roos, B. A., and Signorile, J. F. (2012). Pilot study comparing changes in postural control after training using a video game balance board program and 2 standard activity-based balance intervention programs. *Arch. Phys. Med. Rehabil.* 93, 1138–1146. doi: 10.1016/j.apmr.2012.01.023

Qutubuddin, A. A., Cifu, D. X., Armistead-Jehle, P., Carne, W., McGuirk, T. E., and Baron, M. S. (2007). A comparison of computerized dynamic posturography therapy to standard balance physical therapy in individuals with Parkinson's disease: a pilot study. *NeuroRehabilitation* 22, 261–265.

Rieger, M. M., Papegaaij, S., Pijnappels, M., Steenbrink, F., and van Dieën, J. H. (2020). Transfer and retention effects of gait training with anterior-posterior perturbations to postural responses after medio-lateral gait perturbations in older adults. *Clin. Biomech.* 75, N.PAG-N.PAG. doi: 10.1016/j.clinbiomech.2020.104988

Robinovitch, S. N., Feldman, F., Yang, Y., Schonnop, R., Leung, P. M., Sarraf, T., et al. (2013). Video capture of the circumstances of falls in elderly people residing in long-term care: an observational study. *Lancet.* 381, 47–54. doi: 10.1016/S0140-6736(12)61263-X

Rossi, L. P., Brandalize, M., Pereira, R., and Silveira Gomes, A. R. (2014). The effects of a perturbation-based balance training on neuromuscular recruitment and functional mobility in community-dwelling older women. *Top. Geriatr. Rehabil.* 30, 256–263. doi: 10.1097/TGR.0000000000000035

Rücker, G., Cates, C. J., and Schwarzer, G. (2017). Methods for including information from multi-arm trials in pairwise meta-analysis. *Res. Synth. Method.* 8, 392–403. doi: 10.1002/jrsm.1259

Salanti, G., Giovane, C. D., Chaimani, A., Caldwell, D. M., and Higgins, J. P. T. (2014). Evaluating the quality of evidence from a network meta-analysis. *PLoS ONE.* 9, e99682. doi: 10.1371/journal.pone.0099682

Santos, S. M., da Silva, R. A., Terra, M. B., Almeida, I. A., de Melo, L. B., and Ferraz, H. B. (2017). Balance versus resistance training on postural control in patients with Parkinson's disease: a randomized controlled trial. *Eur. J. Phys. Rehabil. Med.* 53, 173–183. doi: 10.23736/S1973-9087.16.04313-6

Schlenstedt, C., Paschen, S., Kruse, A., Raethjen, J., Weisser, B., and Deuschl, G. (2015). Resistance versus balance training to improve postural control in Parkinson's disease: A randomized rater blinded controlled study. *PLoS ONE* 10. doi: 10.1371/journal.pone.0140584

Schwarzer, G., Carpenter, J. R., and Rücker, G. (2015). Network meta-analysis. In *Meta-Analysis with R* Use R!, eds. *G. Schwarzer, J. R. Carpenter, and G. Rücker (Cham: Springer International Publishing)* p. 187–216. doi: 10.1007/978-3-319-21416-0

Shimada, H., Uchiyama, Y., and Kakurai, S. (2003). Specific effects of balance and gait exercises on physical function among the frail elderly. *Clin. Rehabil.* 17, 472–479. doi: 10.1191/0269215503cr638oa

Shumway-Cook, A., and Woollacott, M. H. (2017). *Motor Control: Translating Research Into Clinical Practice*. Lippincott Williams and Wilkins.

Sibley, K. M., Thomas, S. M., Veroniki, A. A., Rodrigues, M., Hamid, J. S., Lachance, C. C., et al. (2021). Comparative effectiveness of exercise interventions for preventing falls in older adults: a secondary analysis of a systematic review with network meta-analysis. *Exper. Gerontol.* 143, 111151. doi: 10.1016/j.exger.2020.111151

Sohn, J., and Kim, S. (2015). Falls study: Proprioception, postural stability, and slips. *Biomed. Mater. Eng.* 26, S693–703. doi: 10.3233/BME-15 1361

Sterne, J. A. C., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., et al. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ.* 366, l4898. doi: 10.1136/bmj.l4898

Sutton, A. J., Abrams, K. R., Jones, D. R., Sheldon, T. A., and Song, F. (2000). *Methods for Meta-Analysis in Medical Research*. John Wiley and Sons.

Thomas, M., and Kalicinski, M. (2016). The Effects of Slackline Balance Training on Postural Control in Older Adults. *J. Aging Phys. Act.* 24, 393–398. doi: 10.1123/japa.2015-0099

Tinetti, M. E., Speechley, M., and Ginter, S. F. (1988). Risk Factors for Falls among Elderly Persons Living in the Community. *New Engl. J. Med.* 319, 1701–1707. doi: 10.1056/NEJM198812293192604

Wang, Y., Bhatt, T., Liu, X., Wang, S., Lee, A., Wang, E., et al. (2019). Can treadmill-slip perturbation training reduce immediate risk of over-ground-slip induced fall among community-dwelling older adults? *J. Biomech.* 84, 58–66. doi: 10.1016/j.jbiomech.2018.12.017

Winter, D. A., Patla, A. E., and Frank, J. S. (1990). Assessment of balance control in humans. *Med. Prog. Technol.* 16, 31–51.

Wolf, S. L., Barnhart, H. X., Ellison, G. L., Coogler, C. E., and Atlanta, F. I. C. S. I. T., Group (1997). The effect of Tai Chi Quan and computerized balance training on postural stability in older subjects. *Phys. Therapy.* 77, 371–381. doi: 10.1093/ptj/77.4.371

Wooten, S. V., Signorile, J. F., Desai, S. S., Paine, A. K., and Mooney, K. (2018). Yoga meditation (YoMed) and its effect on proprioception and balance function in elders who have fallen: A randomized control study. *Complement Ther. Med.* 36, 129–136. doi: 10.1016/j.ctim.2017.12.010

World Health Organization (2008). *WHO Global Report on Falls Prevention in Older Age*. World Health Organization.

Wu, P.-Y., Huang, K.-S., Chen, K.-M., Chou, C.-P., and Tu, Y.-K. (2021). Exercise, nutrition, and combined exercise and nutrition in older adults with sarcopenia: a systematic review and network meta-analysis. *Maturitas.* 145, 38–48. doi: 10.1016/j.maturitas.2020.12.009

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 11 February 2022
doi: 10.3389/fnagi.2022.838674

Divergence Between Informant and Self-Ratings of Activities of Daily Living Impairments in Parkinson's Disease

*Sara Becker[1,2], Susanne Solbrig[1,3], Katja Michaelis[1,3], Bettina Faust[1,3], Kathrin Brockmann[1,3] and Inga Liepelt-Scarfone[1,3,4]**

[1] Department of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany, [2] Department of Psychology, Faculty of Arts, University of Calgary, Calgary, AB, Canada, [3] German Center for Neurodegenerative Diseases (DZNE), University of Tübingen, Tübingen, Germany, [4] Studienzentrum Stuttgart, IB Hochschule für Gesundheit und Soziales, Stuttgart, Germany

Objective: To examine the agreement between self- and informant-reported activities of daily living (ADL) deficits in Parkinson's Disease (PD) patients, and to examine factors influencing ADL ratings.

Background: In PD, the loss of functional independence is an important outcome of disease progression. The valid assessment of ADL function in PD is essential, but it is unclear to what extent informants' and patients' perceptions of their daily functions concur, and how other factors may influence both ratings.

Methods: Data of 150 PD patients who underwent cognitive and motor testing, as well as their informants were analyzed. The 10-item Functional Activities Questionnaire (FAQ), completed separately by patients (FAQ-S) and their informants (FAQ-I), assessed ADL function. Weighted κ statistics summarized level of agreement, and a discrepancy score (FAQ-I – FAQ-S) quantified agreement. Correlation analyses between FAQ total scores, patient and informant characteristics, and cognitive scores were conducted, with *post hoc* regressions to determine the associations between both FAQ scores and cognition, independent of patient characteristics.

Results: The sample included 87 patients with normal cognition, 50 with mild cognitive impairment, and 13 with dementia. Overall, there was fair to moderate agreement between patients and informants on individual FAQ items ($0.27 \leq \kappa \leq 0.61$, $p < 0.004$), with greater discrepancies with increasing cognitive impairment. Patients' age, motor severity, non-motor burden, and depression also affected both ratings ($0.27 \leq r \leq 0.50$, $p < 0.001$), with motor severity showing the greatest influence on both ratings. Both the FAQ-I and FAQ-S were correlated with almost all cognitive domains. *Post hoc* regression analyses controlling for patient characteristics showed that the attention domain was a significant predictor of both the FAQ-S and FAQ-I scores, and memory was also a significant predictor of the FAQ-I score. Only 29.3% of patients agreed perfectly with informants on the FAQ total score, with informants most commonly rating ADL impairments as more severe than patients.

OPEN ACCESS

Edited by:
Jee-Young Lee, Seoul Metropolitan Government – Seoul National University Boramae Medical Center, South Korea

Reviewed by:
Sanghee Moon, Ithaca College, United States
Chaewon Shin, Chungnam National University, South Korea

****Correspondence:***
Inga Liepelt-Scarfone
inga.liepelt@uni-tuebingen.de

Specialty section:
This article was submitted to Parkinson's Disease and Aging-related Movement Disorders, a section of the journal Frontiers in Aging Neuroscience

Received: *18 December 2021*
Accepted: *21 January 2022*
Published: *11 February 2022*

Citation:
Becker S, Solbrig S, Michaelis K, Faust B, Brockmann K and Liepelt-Scarfone I (2022) Divergence Between Informant and Self-Ratings of Activities of Daily Living Impairments in Parkinson's Disease. Front. Aging Neurosci. 14:838674. doi: 10.3389/fnagi.2022.838674

Conclusions: Patient and informant ratings of ADL function using FAQ items showed moderate agreement, with only few items reaching substantial agreement. Ratings of both were associated with patient cognitive status, but also other characteristics. In addition to patient and informant reports, objective measures are needed to accurately classify ADL deficits in PD.

Keywords: activities of daily living, caregiver, cognition, Functional Activities Questionnaire, Parkinson's Disease, self-ratings, informant-ratings

INTRODUCTION

Significant impairments in activities of daily living (ADL) function, in addition to impaired cognition, are the core criterion for diagnosing Parkinson's Disease (PD) dementia (PDD) (Emre et al., 2007). Both ADL impairments and severe cognitive impairment result in increased risk for nursing home placement and mortality (Hosking et al., 2021). Recent studies have shown that even patients with mild cognitive impairment (PD-MCI) display first signs of ADL dysfunction (Pirogovsky et al., 2014; Cheon et al., 2015; Fellows and Schmitter-Edgecombe, 2019), possibly indicating a group at risk for dementia (Beyle et al., 2018). As the diagnosis of ADL deficits requires these to be solely caused by cognitive deficits, an important step for accurate diagnoses is the measurement of ADL deficits in PD.

Insight into patients' general ADL function is commonly given by a reliable informant, such as a spouse or close friend (Cahn-Weiner et al., 2007). Research in Alzheimer's Disease shows that while informant reporting accurately reflects the functional changes, there are variations in the quality of their reports (Farias et al., 2017). Furthermore, caregiver stress and any depressive symptoms of the caregiver have an influence on the external assessment of ADL in mild dementia patients (Zanetti et al., 1999; Razani et al., 2007). Factors such as caregiver age, education level, living situation, and the nature of the relationship to the patient have been reported to influence the caregivers' assessments in Alzheimer's Disease (Lin et al., 2017), and also in PD (Bhimani, 2014; Mosley et al., 2017; Kalampokini et al., 2020). Additional negative influences on caregivers of PD patients include cognitive status, disease duration, and patients' motor symptom severity (Caap-Ahlgren and Dehlin, 2002; Ransmayr, 2020).

Self- reports are also used to gain perspective on how ADL impairments affect the patient's functioning (Foster and Hershey, 2011). It is important to note that the assessment of ADL impairments in elderly patients becomes more difficult with increasing cognitive deficits; patients with dementia often lack insight to correctly perceive the severity of their illness (Farias et al., 2017). PD patients also tend to underestimate their abilities with increasing cognitive decline (Seltzer et al., 2001), and previous studies have shown PD patients rate themselves as less impaired on measures of ADL than their caregivers (Leritz et al., 2004). Compared to objective measures, 44% of all study participants underestimated their ADL impairment, while 13% overestimated impairment (Shulman et al., 2006). Patients who underestimated ADL disabilities had shorter disease durations, more preserved cognitive abilities, and were living in a family environment, while those who overestimated their ADL skills had advanced PD, showed cognitive dysfunctions, and lived alone (Shulman et al., 2006).

In contrast to the above-mentioned studies, other researchers have not been able to find differences in caregiver versus self-report of ADL disabilities in PD patients (Brown et al., 1989; Liepelt-Scarfone et al., 2013; Copeland et al., 2016). More studies are needed to determine whether self-reports or informant-reports are more useful for judging impairments in ADL function in the clinical routine, as the loss of functional independence is an important outcome of disease progression (Santos Garcia et al., 2021). The aim of this study was therefore to examine both self- and informant-reported ADL using a widely known questionnaire in a cohort of PD patients with varying degrees of cognitive impairment. We aimed to look at the agreement between both sources as well as associations with both patient and informant characteristics, hypothesizing that there would only be moderate agreement between both sources regarding ADL function, with increasing divergence relating to the severity of cognitive impairment.

MATERIALS AND METHODS

Design and Recruitment

Between July 2018 and September 2020, 270 PD patients were recruited to take part in the cross-sectional "Cognitive-driven ADL impairment as a predictor for Parkinson's disease dementia (PDD)" study. Inclusion criteria included: age between 50 and 90 years of age, diagnosis of PD according to UK Brain Bank Criteria, and the ability to understand study requirements and communicate with investigator. Exclusion criteria included: other neurodegenerative disease interfering with cognition or preventing the ability to give informed consent, alcohol, medication or drug dependency or abuse (except for nicotine), or participation in a clinical investigation of a new compound within the last 4 weeks. Additionally, all patients were asked to designate one person to be their informant who was then contacted to give information regarding the patient. This study was approved by the Ethics Committee of the Medical Faculty of the University of Tübingen (284/2018BO1). All patients (or their proxies if necessary) and their informants provided written, informed consent.

Of all patients invited to participate, 36 (13.3%) met exclusion criteria and 52 (19.3%) declined to take part in the examination. A total of 182 (67.4%) patients were included in the study and assessed. For the following final analyses, 17 (9.3%) patients who had received deep brain stimulation (DBS) implantation prior to

assessment were excluded. We chose to exclude these patients as it is currently unclear how DBS affects ADL: some studies have shown improvements on ADL functions (Gorecka-Mazur et al., 2019; Cernera et al., 2020), while others demonstrated no effect on ADL function (Pusswald et al., 2019). Two patients (1.1%) were additionally excluded as medication intake interfered with correct classification of cognitive status. Thirteen (7.1%) patients had missing FAQ data ($n = 9$ did not have an informant and $n = 4$ unable to fill out the FAQ themselves due to severely impaired cognition) and were also excluded. In total, data of 150 patients (and their informants) was analyzed in the final dataset.

Assessments

The Functional Activities Questionnaire (FAQ) (Pfeffer et al., 1982) was used to assess ADL impairments. It consists of 10 items assessing instrumental ADL functions, where capability of each item is rated from 0 (normal) to 3 (dependent on others). A maximum score of 30 points can be achieved, and higher scores indicate greater severity of ADL impairment. The FAQ was separately administered to both the patient as a self-report (FAQ-S) and the informant (FAQ-I) to evaluate the patients' ADL functioning in the previous 4 weeks.

Patient Measures

Patient demographics (age, sex, education years) and medical history [age at onset of PD, disease duration, and anti-parkinsonian medication intake expressed using the levodopa equivalent daily dose (LEDD) (Tomlinson et al., 2010)] were collected. Motor function was assessed by a trained specialist using the Movement Disorder Society Unified Parkinson's Disease Rating Scale Part III (MDS-UPDRS-III) (Goetz et al., 2008). Depression was assessed using the Beck Depression Inventory-II (BDI-II) (Beck et al., 1996), a 21-item instrument quantifying depressive symptoms over the last 2 weeks. Non-motor symptom burden was assessed using the Parkinson's Disease Non-Motor Symptoms Questionnaire (NMSQ) (Chaudhuri et al., 2006).

Global cognitive functioning was measured using the Montreal Cognitive Assessment (MoCA) (Nasreddine et al., 2005). A comprehensive neuropsychological battery assessing five cognitive domains was administered:

- *Attention*: Letter-Number-Sequencing and Digit-Symbol subtests of the Wechsler Intelligenztest für Erwachsene [WIE, German adaptation of the Wechsler Intelligence Test for Adults, (Aster et al., 2006)];
- *Executive Functions*: Trail Making Test-Part B, semantic fluency, and phonemic fluency subtests from the Consortium to Establish a Registry for Alzheimer's Disease–Plus Battery (CERAD-PLUS) (Morris et al., 1988);
- *Memory*: Word List learning, Word List recall, Word List discriminability, and Constructional Praxis recall subtests of the CERAD-PLUS;
- *Visuospatial Functions*: Constructional Praxis subtest of the CERAD-PLUS and Fragmented Words subtest of the "Leistungsprüfsystem 50 + " [performance test for older adults 50–90 years, (Sturm et al., 1993)];
- *Language*: Similarities subtest from the WIE and the modified Boston Naming Test of the CERAD-PLUS.

All raw scores were converted to *z*-scores using test manuals, adjusting for age and/or education where appropriate, and composite domain *z*-scores were calculated for each cognitive domain. Cognitive tests assigned to each domain were chosen according to the recommendations of the Movement Disorders Task Force (Litvan et al., 2012) and this comprehensive battery been previously used in a study with PD patients (Becker et al., 2020). The CERAD-PLUS battery has been shown to be accurate and useful in identifying cognitive impairment in PD patients (Karrasch et al., 2013; Camargo et al., 2018), while both the WIE [English version: Wechsler Adult Intelligence Scale (Yamawaki et al., 2020; Chen et al., 2021)] and the Leistungsprüfsystem 50 + batteries have been utilized in PD-cognition studies (Fengler et al., 2016; Kalbe et al., 2016). Furthermore, a recent systematic literature review identified specific tests used in PD research that have been normed for German-speaking populations and their corresponding cognitive domain (Liepelt-Scarfone et al., 2021). The authors presented guidelines for the neuropsychological assessment of PD patients in the German language, with their findings supporting our chosen tests and domains.

Patients were classified as cognitively normal (PD-CN) if no cognitive or ADL impairment was present. PD-MCI was diagnosed according to the Level-II criteria of the Movement Disorders Task Force (Litvan et al., 2012) if impairment (1.5 standard deviations below population norms) was present in at least two cognitive tests, with preserved ADL functioning. PDD was diagnosed according to consensus criteria (Dubois et al., 2007) if both cognitive impairment and severe impairments in ADL function were present. Fourteen patients ($n = 1$ PD-CN, $n = 4$ PD-MCI, and $n = 9$ PDD) were unable to complete the neuropsychological test battery for various reasons (e.g., severe cognitive impairment, physical and/or mental exhaustion especially toward the end of the test battery). Cognitive diagnoses for these patients were made according to available cognitive data (*z*-scores of the tests the patients did complete), agreements between informants and clinicians regarding ADL status, and neuropsychological investigator judgment. Available medical data (e.g., if the patient had received a diagnosis of either PD-MCI or PDD from a neuropsychologist, neurologist, or primary physician prior to examination) and previous cognitive evaluations were also taken into consideration.

Informant Measures

Demographic information was collected from each informant, including age, sex, education years, living situation, and how many times per week (1× a week, 2–3× per week, or daily) they saw the patient. The Bayer-ADL scale (Hindmarch et al., 1998) was given to informants to assess instrumental ADL, where the patient's ability to perform 25 tasks is rated from 0 "never" to 10 "always." The total sum score is divided by the number of questions answered to obtain a scaled score, ranging from 1 to 10, where higher values indicate more severe impairments in ADL function. Scoring is as follows: 1.0–2.0, no difficulties with ADL,

TABLE 1 | Patient characteristics according to cognitive status.

	PD-CN *n* = 87	PD-MCI *n* = 50	PDD *n* = 13	*p*-value
Demographics				
Male sex: *n* (%)	58 (66.7)	27 (54)	9 (69.2)	0.30
Age (y)	68.45 (51.99–83.47)	68.13 (52.97–83.67)	73.35 (67.69–82.30)	0.17
Education (y)	13 (8–21)	12 (8–19)	12 (11–18)	0.02*
Age at onset (y)	60.48 (39–75.54)	58.77 (44.75–76.58)	64.88 (54.17–76.65)	0.45
Disease duration (y)	7.07 (3.28–21.74)	8.05 (2.15–20.22)	7.42 (5.57–13.92)	0.36
UPDRS-III	22.5 (3–50)	30 (8–68)	45 (27–56)	<0.001**
LEDD	700 (100–1950)	702 (52–1510)	540 (285–1050)	0.45
NMSQ	8 (1–22)	9.5 (0–24)	11 (8–23)	0.03*
BDI-II	8 (0–35)	10 (0–33)	12 (4–27)	0.02*
ADL				
FAQ-I total score	1 (0–24)	3.50 (0–24)	19 (7–28)	<0.001**
FAQ-S total score	0 (0–22)	1 (0–24)	12 (1–28)	<0.001**
Bayer-ADL	1.42 (1–6.12)	2.32 (1–7.20)	6.33 (2.20–9.56)	<0.001**
Cognition‡				<0.001**
MoCA total score	27 (19–30)	24.5 (17–30)	18 (17–19)	<0.001**
Attention	0.20 (−1.20–1.90)	−0.45 (−2.60–1.20)	−1.10 (−2.10−0.70)	<0.001**
Executive functions	0.27 (−2.00–2.57)	−0.83 (–1.93–1.27)	−1.15 (–1.83– −0.70)	<0.001**
Memory	0.29 (−1.90–1.30)	−1.05 (−2.95–0.38)	−1.87 (−2.90– −0.77)	<0.001**
Visuospatial functions	−0.15 (−1.70–1.45)	−1.20 (−2.50–1.00)	−1.62 (−2.65– −1.30)	<0.001**
Language	0.43 (−0.70–1.40)	−0.30 (−1.85–1.35)	−0.47 (−0.70– −0.05)	<0.001**

*Results are given as Median (Range) unless otherwise indicated, *p < 0.05, **p < 0.01. ‡Due to missing values, cognitive domain scores only computed and analyses run for 136 patients. ADL, activities of daily living; BDI-II, Beck Depression Inventory-II; FAQ-I, Functional Activities Questionnaire Informant-rated; FAQ-S, Functional Activities Questionnaire Self-rated; H&Y, Hoehn and Yahr; LEDD, levodopa equivalent daily dose; MoCA, Montreal Cognitive Assessment; NMSQ, Non-Motor Symptoms Questionnaire; PD-CN, Parkinson's Disease cognitively normal; PDD, Parkinson's Disease dementia; PD-MCI, Parkinson's Disease with mild cognitive impairment; UPDRS-III, Unified Parkinson's Disease Rating Scale Part III; y, year.*

2.1–5.0 mild difficulties with everyday function, and 5.1–10.0 indicates clear difficulties in coping with everyday life.

Statistical Analyses

REDCap electronic data capture tools (Harris et al., 2009) hosted at the Hertie Institute for Clinical Brain Research was used to collect and manage study data. SPSS Version 27 (IBM Corp., Armonk, NY, United States) was used to conduct all statistical analyses, with α levels set at 0.05. For missing demographic patient data [n = 12 (8%) MDS-UPDRS-III, n = 4 (2.7%) NMSQ, and n = 2 (1.3%) BDI-II], median values per cognitive group status were imputed to compensate for any missing values. Analyses involving cognitive data only included only those patients who had completed all tests (n = 136) to ensure equal representation of each averaged domain z-score. The Shapiro-Wilk tested assumptions of normality. As data were not normally distributed, demographic variables were examined using the non-parametric Pearson Chi-square, Mann-Whitney U tests, or Independent-Samples Jonckheere-Terpstra Tests for Ordered Alternatives where appropriate.

Weighted κ statistics with linear weights were used to summarize the level of agreement between patient and informants for each individual FAQ item. Agreement values were interpreted as follows: 0–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial, and 0.81–1.00 almost perfect (Landis and Koch, 1977). The FAQ total score was compared between raters using Intraclass Correlation (ICC) using an absolute-agreement, two-way mixed effects model based on single measurements. To determine how different factors may influence the level of agreement, analyses were re-run using the following stratifications: (i) cognitive status of the patient [non-demented (PD-CN and PD-MCI) and demented (PDD)], (ii) sex of the patient, (iii) disease duration using a median split (duration $\leq$ 7.31 years and duration $>$ 7.32 years), and (iv) BDI-II using a median split (score $\leq$ 9 and score $>$ 10).

Spearman's rank correlations between the total FAQ-S and FAQ-I scores, cognitive scores, patient demographic variables of interest (age, sex, education, disease duration, UPDRS-III total score, BDI-II score, and NMSQ score), caregiver variables of interest (age, education years), and the Bayer-ADL were conducted. *Post hoc* multivariate linear regressions were conducted to determine the associations between both FAQ scores and cognition, independent of patient characteristics. Ten regressions were run for each of the five cognitive domains separately (due to multicollinearity when all domains are added into one model), with both the FAQ-S and the FAQ-I as the dependent variable. Further covariates in the models included patient age, UPDRS-III score, NMSQ scores and BDI-II score, as these were all significantly correlated with both FAQ scores. For these regressions, the α was adjusted to 0.005 to correct for multiple comparisons (0.05/10).

Lastly, a discrepancy score (D) was calculated for the total FAQ score between informant and self-ratings: FAQ-I – FAQ-S. Positive values denoted higher impairment rated by informants

(D_I), negative values denoted higher impairment rated by patients themselves (D_S), and scores of 0 indicated perfect agreement between informants and patients (D_A). Patients were split according to the discrepancy score, and demographic and cognitive variables were compared between groups using independent samples Kruskal-Wallis *H* tests, with *post hoc* Bonferroni corrections for multiple testing.

RESULTS

Of all 150 PD patients, 87 (58%) were classified as PD-CN, 50 (33.3%) as PD-MCI, and 13 (8.7%) as PDD. The Jonckheere-Terpstra test showed a significant effect of education between groups, and *post hoc* analyses with Bonferroni correction for multiple tests revealed PD-MCI patients had significantly lower years of education than PD-CN patients ($p = 0.03$; see **Table 1** for details). Significant differences were found between all three groups for the UPDRS-III total score, where PDD patients had the most severe motor impairment according to *post hoc* analyses (PD-CN < PD-MCI < PDD, $p < 0.002$). Analyses also showed PDD patients also had a significantly higher non-motor symptom burden than PD-CN patients (*post hoc* $p = 0.03$). The BDI-II was statistically different between groups, however, *post hoc* significances did not reveal any specific group differences after correction for multiple testing. For measures of ADL, both the FAQ-I and FAQ-S total scores as well as the Bayer-ADL score were significantly different between groups, with PDD patients again showing the most severe impairments in ADL (*post hoc* PD-CN < PD-MCI < PDD, $p < 0.001$).

Comparisons of cognitive data were done using data of 136 patients with complete neuropsychological testing (PD-CN $n = 86$, PD-MCI $n = 46$, PDD $n = 4$). Global cognition measured using the MoCA showed significant differences between groups, with PDD patients exhibiting the most impaired cognitive performance (*post hoc* PD-CN > PD-MCI > PDD, $p < 0.008$). All five cognitive domains were significantly different between groups, following the same *post hoc* pattern where PD-MCI and PDD patients performed similarly, and PD-CN patients performed the best (PD-CN > PD-MCI = PDD, $p < 0.005$).

Regarding informants, they were most frequently spouses (118, 78.7%), followed by children/stepchildren (13, 8.7%), life partner (8, 5.3%), close friend (6, 4%), other relative (3, 2%), and siblings (2, 1.3%) of the patients. Informants were also predominantly females (101, 67.3%), with a median age 64 (range 29–86) and median of 13 years total education (range 8–28). Regarding time spent with the patient, 121 (87.7%) informants reported seeing them daily, 10 (7.2%) two to three times a week, and 7 (5.1%) only once a week.

Agreement Statistics

Table 2 shows the weighted κ statistics for each FAQ item when examining the entire sample. Item 10 (traveling out of house) was rated as substantial agreement ($\kappa = 0.61$, $p < 0.01$), with most other items reaching moderate or fair agreement ($0.27 \leq \kappa \leq 0.59$, $p < 0.004$). A good degree of reliability was found between patient and informant total scores on the FAQ. The single measure ICC was 0.73 with a 95% confidence interval 0.61–0.81 [$F(149,149) = 7.24$, $p < 0.001$]. Splitting patients according to cognitive status revealed that non-demented (PD-CN and PD-MCI) patients and their informants showed fair to moderate agreement on all items ($0.24 \leq \kappa \leq 0.57$, $p < 0.004$), but only 3 items (1, 5, and 10) actually reached the moderate level (see **Table 3**). Agreement between demented patients and their informants was due to chance for almost all items ($0.08 \leq \kappa \leq 0.38$, $p > 0.09$). Only the FAQ items 3 (shopping alone) and 8 (paying attention) demonstrated moderate agreement ($\kappa = 0.56$ and $\kappa = 0.51$, respectively; $p < 0.01$) between patient and informant ratings.

Next, the role of sex of the patient was examined. There was a slight tendency for better agreements when the patient was a male than if they were females (**Supplementary Table 1**). The agreement of ratings for male patients ranged from fair to substantial ($0.30 \leq \kappa \leq 0.68$, $p < 0.002$), while ratings for female patients were fair to moderate ($0.23 \leq \kappa \leq 0.55$, $p < 0.008$) and one item with a slight agreement due to chance ($\kappa = 0.10$, $p = 0.30$). Patients were then split according to median years of disease duration (**Supplementary Table 2**). For patients with a shorter disease duration, item 5 (using appliances) was rated as substantial ($\kappa = 0.62$, $p < 0.001$), with all other items reaching a fair or moderate agreement ($0.22 \leq \kappa \leq 0.53$, $p < 0.01$). The group of patients with longer disease duration also showed fair to moderate agreement ($0.24 \leq \kappa \leq 0.46$, $p < 0.01$), and items 1 (handling finances) and 10 (traveling out of house) reached substantial agreement ($\kappa = 0.62$ and $\kappa = 0.69$, respectively; $p < 0.001$). Lastly, patients were split according to median BDI-II score (**Supplementary Table 3**). Those patients with lower depressive symptomatology demonstrated substantial agreement on three items ($0.62 \leq \kappa \leq 0.72$, $p < 0.001$), while the others showed fair or moderate agreement ($0.31 \leq \kappa \leq 0.52$, $p < 0.001$). Patients with higher depressive symptoms showed overall less agreement than those with higher symptoms. Almost all items had fair to moderate agreement ($0.32 \leq \kappa \leq 0.59$, $p < 0.01$), while item 4 (skills and hobbies) had only slight agreement ($\kappa = 0.17$, $p = 0.02$).

For all agreement analyses run, the number of ratings of items ≥ 1 (indicating at least mild difficulties with the daily task) was consistently higher for informant ratings than for the self-ratings. We examined the consistency of the individual item agreement statistics *post hoc*, to determine whether some FAQ items consistently showed better or worse agreement than others. This was done by ranking the weighted κ statistics of all analyses according to FAQ item and examining the frequency of the items corresponding to the top and bottom three ranks. There was a clear tendency for the items 10 (traveling out of house), 1 (handling finances and balancing checkbook), and a tie between 3 (shopping alone) and 5 (using household appliances) to show the most agreement (top ranked in 8, 7, 4, and 4 analyses, respectively). The items 4 (engaging in skills and hobbies), 6 (preparing a balanced meal), and 7 (keeping up with current events) were those with the poorest consistent agreement (bottom ranked in 8, 6, and 6 analyses, respectively).

Correlation Analyses

The FAQ-S and FAQ-I total scores were moderately positively correlated with one another ($r_s = 0.61$, $p < 0.001$). **Table 4** shows

TABLE 2 | Differences between self and informant ratings of the FAQ.

FAQ item	Self	Informant	Weighted κ	SE of κ	*p*-value
1. Handling finances	20 (13.3)	36 (24)	0.59	0.08	<0.001**
2. Assembling tax records	47 (31.3)	56 (37.3)	0.47	0.07	<0.001**
3. Shopping	27 (18)	41 (27.3)	0.48	0.08	<0.001**
4. Skills and hobbies	21 (14)	58 (38.7)	0.27	0.06	<0.001**
5. Using appliances	12 (8)	24 (16)	0.48	0.11	<0.001**
6. Meal preparation	31 (20.7)	52 (34.7)	0.37	0.08	<0.001**
7. Current events	13 (8.7)	26 (17.3)	0.37	0.10	<0.001**
8. Paying attention	20 (13.3)	30 (20)	0.40	0.08	<0.001**
9. Remembering appointments	37 (24.7)	48 (32)	0.39	0.07	<0.001**
10. Traveling out of house	26 (17.3)	42 (28)	0.61	0.07	<0.001**

*Results are expressed as Number of patients scoring ≥ 1 (%), **p < 0.01. FAQ, Functional Activities Questionnaire; SE, standard error.*

TABLE 3 | Agreement between self and informant ratings of the FAQ according to cognitive status.

	Non-demented PD patients *n = 137*					Demented PD patients *n = 13*				
FAQ item	Self	Informant	Weighted κ	SE of κ	*p*-value	Self	Informant	Weighted κ	SE of κ	*p*-value
1. Handling finances	11 (8)	23 (16.8)	0.51	0.11	<0.001**	9 (69.2)	13 (100)	0.17	0.12	0.18
2. Assembling tax records	37 (27)	43 (31.4)	0.37	0.08	<0.001**	10 (76.9)	13 (100)	0.08	0.07	0.26
3. Shopping	19 (13.9)	30 (21.9)	0.38	0.09	<0.001**	8 (61.5)	11 (84.6)	0.56	0.17	0.002**
4. Skills and hobbies	14 (10.2)	50 (36.5)	0.24	0.07	<0.001**	7 (53.8)	8 (61.5)	0.13	0.18	0.52
5. Using appliances	7 (5.1)	15 (10.9)	0.45	0.15	<0.001**	5 (38.5)	9 (69.2)	0.29	0.20	0.15
6. Meal preparation	24 (17.5)	41 (29.9)	0.29	0.08	<0.001**	7 (53.8)	11 (84.6)	0.21	0.19	0.28
7. Current events	10 (7.3)	18 (13.1)	0.39	0.11	<0.001**	3 (23.1)	8 (61.5)	0.18	0.22	0.33
8. Paying attention	15 (10.9)	24 (17.5)	0.33	0.09	<0.001**	5 (38.5)	6 (46.2)	0.51	0.14	0.01*
9. Remembering appointments	26 (19)	37 (27)	0.30	0.09	<0.001**	11 (84.6)	11 (84.6)	0.20	0.20	0.27
10. Traveling out of house	19 (13.9)	32 (23.4)	0.57	0.09	<0.001**	7 (53.8)	10 (76.9)	0.38	0.19	0.09

*Results are expressed as Number of patients scoring ≥ 1 (%), *p < 0.05, **p < 0.01. FAQ, Functional Activities Questionnaire; PD, Parkinson's Disease; SE, standard error.*

the correlations for the FAQ scores and both patient and caregiver variables of interest. Both the FAQ-S and FAQ-I total scores were positively correlated with patients' age, motor severity, number of non-motor symptoms, and depressive symptoms ($0.27 \leq r_s \leq 0.50$, $p < 0.001$), and negatively correlated with patients' education ($r_s = -0.21$ and $r_s = -0.18$, respectively; $p < 0.03$). The Bayer-ADL score was strongly positively correlated with the FAQ-I ($r_s = 0.85$, $p < 0.001$) and moderately correlated with the FAQ-S ($r_s = 0.55$, $p < 0.001$). Examining the cognitive data, both the FAQ-S and FAQ-I scores were negatively associated with the MoCA ($r_s = -0.33$ and $r_s = -0.38$, respectively; $p < 0.001$). The FAQ-S was significantly negatively correlated with the attention, executive functions, memory, and visuospatial functions domains ($-0.40 \leq r_s \leq -0.19$, $p < 0.02$), but not with language. In contrast, for the FAQ-I, there was a significant negative correlation with all cognitive domains ($-0.43 \leq r_s \leq -0.21$, $p < 0.02$).

Post hoc Regression Analyses

All ten linear regression models with either the FAQ-S or FAQ-I as the dependent variable were stable. The linear regression with attention as the cognitive independent and FAQ-S as the dependent variable was statistically significant [$F(5,130) = 8.27$, $p < 0.001$], explaining 24.1% (Nagelkerke R^2) of the variance. Attention significantly predicted the FAQ-S (unstandardized $\beta = -1.49$, standard error of $\beta = 0.46$, $p = 0.001$), as did the UPDRS-III score (unstandardized $\beta = 0.09$, standard error of $\beta = 0.03$, $p = 0.004$). The model including memory was also significant [$F(5,130) = 7.50$, $p < 0.001$], explaining 22.4% (Nagelkerke R^2) of the variance. Memory domain as a predictor did not reach clinical significance after correction for multiple testing (unstandardized $\beta = -0.99$, standard error of $\beta = 0.36$, $p = 0.007$), while the UPDRS-III score was again a significant predictor (unstandardized $\beta = 0.11$, standard error of $\beta = 0.04$, $p = 0.001$). For the FAQ-S models including executive functions, visuospatial functions, and language domains, only the UPDRS-III score was a significant predictor ($p < 0.004$).

When the dependent variable was the FAQ-I, the model including attention was again significant [$F(5,130) = 10.48$, $p < 0.001$], explaining 28.7% (Nagelkerke R^2) of the variance. Attention significantly predicted the FAQ-I (unstandardized $\beta = -1.75$, standard error of $\beta = 0.56$, $p = 0.002$), as did the UPDRS-III score (unstandardized $\beta = 0.11$, standard error of $\beta = 0.04$, $p = 0.005$) and patient age (unstandardized $\beta = 0.18$, standard error of $\beta = 0.05$, $p < 0.001$). When memory was the independent variable, the model was again significant [$F(5,130) = 10.12$, $p < 0.001$], explaining 28% (Nagelkerke R^2) of the variance. Significant predictors of the FAQ-I were memory (unstandardized $\beta = -1.27$, standard error of $\beta = 0.44$, $p = 0.004$), UPDRS-III score (unstandardized $\beta = 0.12$, standard

TABLE 4 | Correlations between the total FAQ-S and FAQ-I scores and both patient and caregiver variables of interest.

	FAQ-S	FAQ-I
Patient characteristics		
Patient age	0.27**	0.34**
Patient education	−0.21**	−0.18*
Disease duration	0.05	0.13
UPDRS-III	0.38**	0.50**
NMSQ	0.35**	0.36**
BDI-II	0.27**	0.35**
Patient cognition‡		
MoCA total score	−0.33**	−0.38**
Attention	−0.38**	−0.43**
Executive functions	−0.31**	−0.36**
Memory	−0.40**	−0.31**
Visuospatial functions	−0.19*	−0.21*
Language	−0.14	−0.21*
Caregiver characteristics		
Caregiver age	0.06	0.11
Caregiver education	−0.18*	−0.14
Bayer-ADL	0.55**	0.85**

*$p < 0.05$, **$p < 0.01$. ‡Due to missing values, cognitive domain scores only computed and analyses run for 136 patients. ADL, activities of daily living; BDI-II, Beck Depression Inventory-II; FAQ-I, Functional Activities Questionnaire Informant-rated; FAQ-S, Functional Activities Questionnaire Self-rated; MoCA, Montreal Cognitive Assessment; NMSQ, Non-Motor Symptoms Questionnaire; UPDRS-III, Unified Parkinson's Disease Rating Scale Part III.*

error of β = 0.04, $p = 0.002$), and patient age (unstandardized β = 0.19, standard error of β = 0.05, $p < 0.001$). The FAQ-S models including executive functions, visuospatial functions, and language domains, were significant, however, the only significant predictors were patient age ($p < 0.001$) and UPDRS-III score ($p < 0.001$).

Discrepancy Scores

The discrepancy score based on difference between the FAQ-S and FAQ-I total scores showed that 76 (50.7%) of informants rated ADL function as more impaired than their patients (D_I group), 44 (29.3%) of participants agreed with their informants (D_A group), and 30 (20%) patients rated their impairment worse than informants (D_S group). Cognitive group distribution was as follows: D_I – 38 PD-CN, 27 PD-MCI, 11 PDD; D_A – 34 PD-CN, 10 PD-MCI; D_S – 15 PD-CN, 13 PD-MCI, 2 PDD.

Kruskal-Wallis H tests showed group differences for patient age [$H(2) = 6.59$, $p = 0.04$], and *post hoc* Bonferroni corrected analyses revealed patients in the D_I group were significantly older than patients in the D_A group ($p = 0.04$). A significant effect was found for the MDS-UPDRS-III score [$H(2) = 19.29$, $p < 0.001$], where motor severity was greater in the D_I group than in the D_A group ($p < 0.001$). The NMSQ was also significantly different between groups [$H(2) = 11.94$, $p = 0.003$], with patients in the D_A group reporting less non-motor symptoms than both the D_S ($p = 0.03$) and D_I ($p = 0.003$) groups. Depressive symptomatology was also different between groups [$H(2) = 11.29$, $p = 0.004$], where patients in the D_I group had higher BDI scores than those in the D_A group. The Bayer ADL scale score was significantly different between groups, [$H(2) = 57.82$, $p < 0.001$], with *post hoc* analyses showing $D_A < D_S < D_I$, $p < 0.02$ for all. Neither caregiver age [$H(2) = 0.66$, $p = 0.72$] nor education level [$H(2) = 2.15$, $p = 0.34$] were significantly different between groups.

For the cognitive data, a significant effect for the MoCA was found [$H(2) = 9.47$, $p = 0.009$], where patients in the DI group had significantly lowered global cognition than patients in the DA group ($p = 0.006$). The attention domain showed a significant group effect [$H(2) = 22.49$, $p < 0.001$] where patients in the D_A had significantly better cognition than patients in both the D_S and D_I groups ($p < 0.01$). Significant group effects were also found for the executive functions [$H(2) = 13.45$, $p = 0.001$], memory [$H(2) = 9.85$, $p = 0.007$], and visuospatial functions [$H(2) = 9.68$, $p = 0.008$] domains. *Post hoc* results showed patients in the D_A had significantly better cognition than patients in the D_I groups ($p < 0.01$). For the visuospatial domain, patients in the D_A had borderline insignificant ($p = 0.05$) better cognitive scores than patients in the D_S group. No effect was found between groups for performance in the language domain [$H(2) = 1.62$, $p = 0.45$]. **Figure 1** shows a box-and-whisker plot of performance on each cognitive domain across discrepancy score groups including *post hoc* significant differences.

DISCUSSION

This study aimed to examine the agreement between self- and informant-reports of ADL function and determine whether they were affected by different patient or informant characteristics. The main results were that: (i) agreement analyses showed overall fair to moderate agreement, with few items reaching substantial agreement in various sub-analyses; (ii) both patient and informant ratings were significantly correlated with patient characteristics including cognition, as well as motor severity, age, and depression; and (iii) informants most commonly rated ADL impairments as more severe than patients, with only 29.3% of patients showing perfect agreement with informants on the their ADL performance as reflected by the FAQ total score.

Severe deficits in daily functioning are the core criteria for diagnosing PDD, however, studies have shown that even PD-MCI patients can show early, non-clinical signs of ADL dysfunction (Pirogovsky et al., 2014; Beyle et al., 2018). It is important to understand how and when changes in ADL indicative of dementia in PD emerge to be able to provide early interventions for maintaining patients' autonomy and, as a direct consequence, their quality of life. However, to be able to recognize significant changes in ADL indicating first signs of PDD, reliable and valid assessments with high diagnostic accuracy are needed. Our current results showed that when using the FAQ, there was only a fair to moderate agreement between patients and their informants regarding the daily functional abilities of the patient. The data also support previous findings that there was some level of over- and underreporting done by either patients or caregivers (Shulman et al., 2006; Christ et al., 2013; Cholerton et al., 2020), which is important to take into account when using these questionnaires to screen for ADL impairment indicative

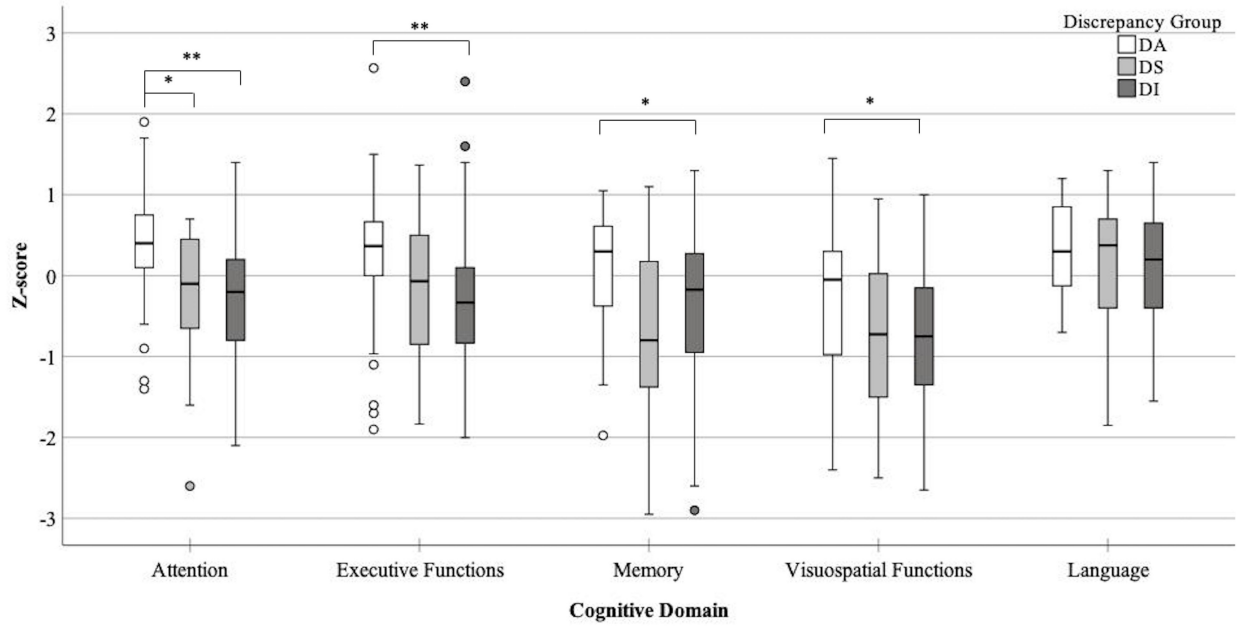


FIGURE 1 | Box-and-whisker plot detailing performance across all five cognitive domains for each discrepancy score group (DA- perfect agreement between informants and patients, DS- higher impairment rated by patients, DI- higher impairment rated by informants), $*p < 0.05$, $**p < 0.001$.

of dementia. To the authors knowledge, only one other study has directly examined inter-rater agreement of ADL ratings by giving both patients and informants the same scale (Deck et al., 2019). Patient and caregiver agreement was examined with the Penn Parkinson's Daily Activities Questionnaire (PDAQ-15), a scale designed to assess instrumental cognitive-associated ADL impairment in PD, and additionally compared to an objective functional measure of ADL. The authors of this study found a moderate agreement between patient and informant ratings on the total (ICC = 0.57) and individual items of the PDAQ-15 (Deck et al., 2019), similar to our current findings reported for the FAQ. Together these results show that agreement between patients and informants is not perfect, regardless of the scale used and the underlying constructs measured. It is imperative to for clinicians and researchers alike to consider that using a single measure of ADL dysfunction, whether rated by the patient or informant, to detect and diagnose presence of PDD potentially overestimates the motor effects on ADL function as well as underestimating cognitive sources. Incorporating objective measures of functional performance as well as self or informant-reported questionnaires may aid in determining level of ADL functioning especially in PD-MCI patients who would be able to undergo rigorous testing. Alternatively, consensus rating between both patient and informants should be considered, although studies are needed to determine whether these will better correspond to patient's real-world abilities than individual assessments.

A specific item analysis was undertaken which showed that traveling out of house, handling finances, going shopping alone, and using household appliances had the highest agreement in our sample, while engaging in skills and hobbies, preparing a balanced meal, and keeping up with current events had the poorest agreement. A previous study showed the specific ADL items managing finances, keeping their appointments, following current events, and using a phone were unaffected by motor symptoms and able to identify dementia in PD patients (Cheon et al., 2015). Furthermore, differentiating cognitive from motor influences on the FAQ showed that both handling finances and keeping up with current events were predicted by cognitive, not motor, abilities (Becker et al., 2020). However, only finances showed adequate agreement within our sample, whereas the agreements between patients and informants regarding following current events were consistently poor. It is possible that impairments in handling finances can be better acknowledged by patients because there is a tangible result which may be cause for concern when impairments are noted, while keeping up with current events is a more subjective experience. Patients may tend to overestimate this ability as they lack a concrete way to measure its loss. On the other hand, traveling out of the house, going shopping alone, and using appliances were all related to motor ADL – as these are affected by PD motor symptoms (e.g., tremor or dyskinesias), it is possible that these are more often noticed by patients and informants alike. This is also supported by the fact that motor severity was a significant factor in predicting both FAQ-S and FAQ-I scores. Patients may be more willing to admit their difficulties in these areas as opposed to, for example, engaging in their skills and hobbies which they are more hesitant to give up due to motor symptoms. Future studies should examine in further detail whether there are specific ADL abilities that may be more prone to disagreements and how both patient and informant perceptions of these abilities can change over time. It is possible that either patient or informant ratings on certain ADL questionnaire items correspond accurately

to patients' real-life functioning. Such analyses can inform judgments regarding severity of ADL deficits as more weight can be given to items and ratings known to reflect patients' daily life.

Generally, studies have focused on comparing informant-rated and patient-rated ADL functioning to the objective cognitive performance of the patient (Shulman et al., 2006; Copeland et al., 2016; Cholerton et al., 2020). While both patients and their informants may not accurately identify specific cognitive deficits, these studies generally show that in early stages of cognitive decline, patient-ratings may be more sensitive to changes in ADL affected by cognition than those of their informants. Only with increasing cognitive decline does the participant lose awareness of their ADL abilities and the knowledgeable informant report become more valuable. Current results confirmed this previous research that PD patients underestimate their abilities with increasing cognitive decline by rating themselves as less impaired than their informants (Seltzer et al., 2001; Leritz et al., 2004; Shulman et al., 2006; Deck et al., 2019). There was poor agreement which was due to chance for almost all individual FAQ items in PDD patients, indicating that demented patients and their informants cannot agree on the patients' ability to carry out ADL. Due to the small sample size in our study, however, data needed to be interpreted with caution and validated in larger samples. When examining non-demented patients, agreement ratings tended to be worse than when examining the entire sample. This is interesting, as anosognosia is not necessarily found in earlier stages of cognitive decline, although reports are varied (Pennington et al., 2021). We also examined associations directly with cognitive domains. Both patient and informant ratings were correlated with almost all cognitive domains, showing that ADL impairments present in PD are again associated with cognitive decline. To determine whether and to what extent this effect was influenced by patient characteristics, we conducted additional regression analyses. These showed that both patient and informant FAQ ratings were predicted by patients' performance on attention and memory domains. Previously, it has been shown that deficits in attention predicted ADL performance in PD, after controlling for confounders including age, sex, and motor impairment (Bronnick et al., 2006; Becker et al., 2020). Our results that attentional deficits, independent of motor or non-motor symptoms, increased ratings of ADL impairment corroborate previous results and highlight that attentional deficits may decline in parallel with ADL deficits. Studies have also demonstrated relationships between memory performance and ADL impairments (Beyle et al., 2018; Foster and Doty, 2021). Future research should confirm how memory and attention deficits impair patients' ADL function, which may lead to more person-centered interventions and assessment strategies.

Apart from cognition, both patient and informant ratings were influenced by patient characteristics, evidenced by both correlation and agreement analyses performed. Severity of PD symptoms, presence of depression, and disease-related motor complications in quality of life are associated with self-assessment of ADL function (Hobson et al., 2001). Our findings replicate most of these findings and also demonstrate they are associated with informant assessments, by showing that both the FAQ-I and FAQ-S were correlated with patients' age, motor severity, non-motor symptom burden, and depression symptoms. This is important as more than 90% of PD patients present with at least one NMS (Barone et al., 2009) and around 35% present with clinically significant depressive symptoms (Reijnders et al., 2008). Clinicians evaluating ADL deficits should ask whether and to what extent these symptoms are affecting daily functioning, to ascertain a more reliable index of ADL function. Perhaps the most important finding is that patients' motor severity influences all ADL ratings, and even modulates the association between cognition and ADL ratings. Previous studies have shown that motor symptoms of PD affect ADL (Martinez-Martin et al., 2003; Skinner et al., 2015) and are independent of the cognitive-driven ADL aspects (Becker et al., 2020). However, the diagnosis of PDD requires ADL deficits to be caused by cognitive and not motor problems; as both patient and informant reports are influenced by motor severity, this further emphasizes the need for accurate ADL scales that can capture deficits independent of motor influences. Furthermore, we found an effect of the patients' sex, where there was a slight tendency for better agreement for male than female patients. This is in line with a previous study that found that women with PD were likely to report more severe ADL dysfunction (Medijainen et al., 2015). This is an interesting finding that should be replicated in future studies, as it would be crucial for clinicians to consider sex differences when judging ADL deficits. Notably, neither informant nor self-ratings were influenced by caregiver age or education, which is not in line with previous literature (Bhimani, 2014; Kalampokini et al., 2020). However, we did not evaluate more specific informant details, such as caregiver burden, depression, or social life, all of which have been shown to have an effect on the reporting of ADL function. Future studies are needed to determine more specifically how and to what extent these caregiver attributes affect their ratings, and whether there are ways to partialize these effects out.

Lastly, discrepancy scores were used to quantify the agreement between patients and informants. Generally, informants rated the ADL impairments as more severe than the patients did, and only 30% of patients and informants agreed regarding the total score. Notably, there were no PDD patients where there was an agreement between both raters, further confirming that PD patients show worse insight into their ADL deficits with worsening cognition. We found that the group where informants rated ADL deficits as more severe were older, had worse motor severity, higher non-motor burden, more depressive symptoms, and worse cognition (in the attention, executive function, memory, and visuospatial domains) than the group where patients and informants agreed. A previous study found patients who underestimated ADL disabilities had shorter disease durations, more preserved cognitive abilities, and were living in a family environment, while those who overestimated their ADL skills had advanced PD, showed cognitive dysfunctions, and lived alone (Shulman et al., 2006). It would be interesting for future studies to determine to what extent difficulties in ADL are results of these factors and the disease progression. Furthermore, studies should explore different factors related to either the patient (such as comorbid

diseases, neuropsychiatric disturbances, or other common non-motor symptoms including fatigue/sleep disturbances, or urinary dysfunction) or the informant (personality traits, quality of life, stress, and psychosocial burden) that could influence ADL function, and how these may affect ratings.

There are limitations of this study that need to be addressed. Most importantly, no objective measures of ADL function, such as performance-based tests which were used in Deck et al. (2019), were used to validate the ADL ratings of the FAQ in this study. Therefore, we cannot determine whether patients' or their informants' ratings correspond most accurately to the real-life functional performance. While performance-based tests have been used in PD research, these measures are often not feasible in clinical and research settings as they are time-consuming and can be heavily influenced by motor severity (Sulzer et al., 2020). Future studies should incorporate objective measures of ADL function to determine the relation of both patient and informant ratings to objective observed ADL performance and to overcome any methodological biases or inaccuracy errors that may be related to the use of questionnaires (Sadek et al., 2011). Moreover, the current analyses should be repeated using different ADL scales, to determine whether they show better or worse agreement between patients and informants, to aid both clinicians and researchers in their selection of assessment tools. The current results should also be interpreted with caution due to the imbalanced group sizes for each cognitive status (especially for the PDD group). Studies should seek to replicate current findings using not only equal sample sizes, but also using different neuropsychological tests or different cut-offs for diagnosis of cognitive impairment. Furthermore, as briefly mentioned, we did not evaluate more specific caregiver details, as the primary outcome of the study was changes in ADL functioning of patients (with information from informants supporting the evaluation of ADL function). Certain sub-analyses of caregiver variables also could not be performed in this study. Examination of whether relation to and time spent with the patient affected agreement were not possible due to a too-small group sizes for adequately powered analyses, and agreement statistics between different categories of caregiver relationship were unable to be run in certain cases due to variables being constant. More studies are needed to determine the extent that informant factors may have on their perceptions of patient ADL functions.

CONCLUSION

This study examined self- and informant-reported ADL using a widely known questionnaire in a cohort of PD patients with varying degrees of cognitive impairment by looking at the agreement between both raters as well as associations with both patient and informant characteristics. Results of the study showed that when using the same ADL questionnaire, there was only a fair to moderate agreement between patients and their informants regarding the patients' daily functioning, with divergence between ratings increasing as cognitive impairment becomes more severe. Overall, less than one-third (29.3%) of PD patients had perfect agreement with informants on the FAQ total score, highlighting again that this discrepancy is very pronounced. Our results also highlight that motor severity influences ADL ratings and modulates associations to cognition, Lastly, as the diagnosis of PDD necessitates impairments in ADL to be cognitive based, it is important for clinicians to understand that both informants and patients are affected by patient characteristics (most notably motor severity) and must take these into account and rule out their influences when deciding on a dementia diagnosis.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the corresponding author will consider any requests for access to the data (including de-identified participant data and corresponding data dictionary) reported in this manuscript. Requests to access the datasets should be directed to IL-S, inga.liepelt@uni-tuebingen.de.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Medical Faculty of University of Tübingen. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

SB, BF, and IL-S: design and conceptualization of study. SB, SS, KM, and KB: data collection. SB and IL-S: statistical analysis. SB: writing of the first draft. SS, KM, BF, KB, and IL-S: review and critique of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by a grant from the Michael J. Fox Foundation (MJFF) for Parkinson's Research (Grant Number 15227). Michael J. Fox Foundation was not involved in the study design, data collection, data analysis, decision to publish, or preparation of the manuscript. We also acknowledge support by the Open Access Publishing Fund of the University of Tübingen.

ACKNOWLEDGMENTS

We would like to thank all patients and informants who participated in this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.838674/full#supplementary-material

REFERENCES

Aster, M., Neubauer, A., and Horn, R. (2006). *WIE. Wechsler Intelligenztest Fuer Erwachsene. Deutschsprachige Bearbeitung und Adaptation des WAIS-III von David Wechsler*. Bern: Hogrefe.

Barone, P., Antonini, A., Colosimo, C., Marconi, R., Morgante, L., Avarello, T. P., et al. (2009). The PRIAMO study: a multicenter assessment of nonmotor symptoms and their impact on quality of life in Parkinson's disease. *Mov. Disord.* 24, 1641–1649. doi: 10.1002/mds.22643

Beck, A. T., Steer, R. A., and Brown, G. K. (1996). *BDI-II, Beck depression inventory : Manual. San Antonio, Tex.; Boston: Psychological Corp; Harcourt Brace*. New York, NY: Springer.

Becker, S., Baumer, A., Maetzler, W., Nussbaum, S., Timmers, M., Van Nueten, L., et al. (2020). Assessment of cognitive-driven activity of daily living impairment in non-demented Parkinson's patients. *J. Neuropsychol.* 14, 69–84. doi: 10.1111/jnp.12173

Beyle, A., Glonnegger, H., Cerff, B., Graber, S., Berg, D., and Liepelt-Scarfone, I. (2018). The Multiple Object Test as a performance-based tool to assess the decline of ADL function in Parkinson's disease. *PLoS One* 13:e0200990. doi: 10.1371/journal.pone.0200990

Bhimani, R. (2014). Understanding the Burden on Caregivers of People with Parkinson's:a Scoping Review of the Literature. *Rehabil. Res. Pract.* 2014:718527. doi: 10.1155/2014/718527

Bronnick, K., Ehrt, U., Emre, M., De Deyn, P. P., Wesnes, K., Tekin, S., et al. (2006). Attentional deficits affect activities of daily living in dementia-associated with Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 77, 1136–1142. doi: 10.1136/jnnp.2006.093146

Brown, R. G., MacCarthy, B., Jahanshahi, M., and Marsden, C. D. (1989). Accuracy of self-reported disability in patients with parkinsonism. *Arch. Neurol.* 46, 955–959. doi: 10.1001/archneur.1989.00520450025014

Caap-Ahlgren, M., and Dehlin, O. (2002). Factors of importance to the caregiver burden experienced by family caregivers of Parkinson's disease patients. *Aging Clin. Exp. Res.* 14, 371–377. doi: 10.1007/BF03324464

Cahn-Weiner, D. A., Farias, S. T., Julian, L., Harvey, D. J., Kramer, J. H., Reed, B. R., et al. (2007). Cognitive and neuroimaging predictors of instrumental activities of daily living. *J. Int. Neuropsychol. Soc.* 13, 747–757. doi: 10.1017/S1355617707070853

Camargo, C. H. F., Bronzini, A., Tolentino, E. S., Medyk, C., and Schultz-Pereira, G. L. (2018). Can the CERAD neuropsychological battery be used to assess cognitive impairment in Parkinson's disease? *Arq. Neuropsiquiatr.* 76, 145–149. doi: 10.1590/0004-282x20180003

Cernera, S., Eisinger, R. S., Wong, J. K., Ho, K. W. D., Lopes, J. L., To, K., et al. (2020). Long-term Parkinson's disease quality of life after staged DBS: STN vs GPi and first vs second lead. *NPJ Park. Dis.* 6:13. doi: 10.1038/s41531-020-0115-3

Chaudhuri, K. R., Martinez-Martin, P., Schapira, A. H., Stocchi, F., Sethi, K., Odin, P., et al. (2006). International multicenter pilot study of the first comprehensive self-completed nonmotor symptoms questionnaire for Parkinson's disease: the NMSQuest study. *Mov. Disord.* 21, 916–923. doi: 10.1002/mds.20844

Chen, Y. S., Chen, H. L., Lu, C. H., Lee, C. Y., Chou, K. H., Chen, M. H., et al. (2021). The corticolimbic structural covariance network as an early predictive biosignature for cognitive impairment in Parkinson's disease. *Sci. Rep.* 11:862. doi: 10.1038/s41598-020-79403-x

Cheon, S. M., Park, K. W., and Kim, J. W. (2015). Identification of Daily Activity Impairments in the Diagnosis of Parkinson Disease Dementia. *Cogn. Behav. Neurol.* 28, 220–228. doi: 10.1097/WNN.0000000000000081

Cholerton, B., Poston, K. L., Tian, L., Quinn, J. F., Chung, K. A., Hiller, A. L., et al. (2020). Participant and Study Partner Reported Impact of Cognition on Functional Activities in Parkinson's Disease. *Mov. Disord. Clin. Pract.* 7, 61–69. doi: 10.1002/mdc3.12870

Christ, J. B., Fruhmann Berger, M., Riedl, E., Prakash, D., Csoti, I., Molt, W., et al. (2013). How precise are activities of daily living scales for the diagnosis of Parkinson's disease dementia? A pilot study. *Park. Relat. Disord.* 19, 371–374. doi: 10.1016/j.parkreldis.2012.11.004

Copeland, J. N., Lieberman, A., Oravivattanakul, S., and Troster, A. I. (2016). Accuracy of Patient and Care Partner Identification of Cognitive Impairments in Parkinson's Disease-Mild Cognitive Impairment. *Mov. Disord.* 31, 693–698. doi: 10.1002/mds.26619

Deck, B. L., Xie, S. X., Choi, G., Rick, J., Siderowf, A., Rudovsky, S., et al. (2019). Cognitive Functional Abilities in Parkinson's Disease:agreement Between Patients and Informants. *Mov. Disord. Clin. Pract.* 6, 440–445. doi: 10.1002/mdc3.12781

Dubois, B., Burn, D., Goetz, C., Aarsland, D., Brown, R. G., Broe, G. A., et al. (2007). Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. *Mov. Disord.* 22, 2314–2324. doi: 10.1002/mds.21844

Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., et al. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov. Disord.* 22, 1689–1707. doi: 10.1002/mds.21507

Farias, S. T., Lau, K., Harvey, D., Denny, K. G., Barba, C., and Mefford, A. N. (2017). Early Functional Limitations in Cognitively Normal Older Adults Predict Diagnostic Conversion to Mild Cognitive Impairment. *J. Am. Geriatr. Soc.* 65, 1152–1158. doi: 10.1111/jgs.14835

Fellows, R. P., and Schmitter-Edgecombe, M. (2019). Multimethod assessment of everyday functioning and memory abilities in Parkinson's disease. *Neuropsychology* 33, 169–177. doi: 10.1037/neu0000505

Fengler, S., Kessler, J., Timmermann, L., Zapf, A., Elben, S., Wojtecki, L., et al. (2016). Screening for Cognitive Impairment in Parkinson's Disease: improving the Diagnostic Utility of the MoCA through Subtest Weighting. *PLoS One* 11:e0159318. doi: 10.1371/journal.pone.0159318

Foster, E. R., and Doty, T. (2021). Cognitive Correlates of Instrumental Activities of Daily Living Performance in Parkinson Disease Without Dementia. *Arch. Rehabil. Res. Clin. Transl.* 3:100138. doi: 10.1016/j.arrct.2021.100138

Foster, E. R., and Hershey, T. (2011). Everyday Executive Function Is Associated With Activity Participation in Parkinson Disease Without Dementia. *OTJR* 31, 16–22. doi: 10.3928/15394492-20101108-04

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., et al. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov. Disord.* 23, 2129–2170. doi: 10.1002/mds.22340

Gorecka-Mazur, A., Furgala, A., Krygowska-Wajs, A., Pietraszko, W., Kwinta, B., and Gil, K. (2019). Activities of Daily Living and Their Relationship to Health-Related Quality of Life in Patients with Parkinson Disease After Subthalamic Nucleus Deep Brain Stimulation. *World Neurosurg.* 125, e552–e562. doi: 10.1016/j.wneu.2019.01.132

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., and Conde, J. G. (2009). Research electronic data capture (REDCap)–a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* 42, 377–381. doi: 10.1016/j.jbi.2008.08.010

Hindmarch, I., Lehfeld, H., de Jongh, P., and Erzigkeit, H. (1998). The Bayer Activities of Daily Living Scale (B-ADL). *Dement. Geriatr. Cogn. Disord.* 9, 20–26. doi: 10.1159/000051195

Hobson, J. P., Edwards, N. I., and Meara, R. J. (2001). The Parkinson's Disease Activities of Daily Living Scale: a new simple and brief subjective measure of disability in Parkinson's disease. *Clin. Rehabil.* 15, 241–246. doi: 10.1191/026921501666767060

Hosking, A., Hommel, A., Lorenzl, S., Coelho, M., Ferreira, J. J., Meissner, W. G., et al. (2021). Characteristics of Patients with Late-Stage Parkinsonism Who are Nursing Home Residents Compared with those Living at Home. *J. Am. Med. Dir. Assoc.* 22, 440–445.e2. doi: 10.1016/j.jamda.2020.06.025

Kalampokini, S., Hommel, A., Lorenzl, S., Ferreira, J. J., Meissner, W. G., Odin, P., et al. (2020). Caregiver Burden in Late-Stage Parkinsonism and Its Associations. *J. Geriatr. Psychiatry. Neurol.* 35, 110–120. doi: 10.1177/0891988720968263

Kalbe, E., Rehberg, S. P., Heber, I., Kronenbuerger, M., Schulz, J. B., Storch, A., et al. (2016). Subtypes of mild cognitive impairment in patients with Parkinson's disease: evidence from the LANDSCAPE study. *J. Neurol. Neurosurg. Psychiatry* 87, 1099–1105. doi: 10.1136/jnnp-2016-313838

Karrasch, M., Laatu, S., Martikainen, K., and Marttila, R. (2013). CERAD test performance and cognitive impairment in Parkinson's disease. *Acta Neurol. Scand.* 128, 409–413. doi: 10.1111/ane.12138

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174. doi: 10.2307/2529310

Leritz, E., Loftis, C., Crucian, G., Friedman, W., and Bowers, D. (2004). Self-awareness of deficits in Parkinson disease. *Clin. Neuropsychol.* 18, 352–361. doi: 10.1080/1385404049052412

Liepelt-Scarfone, I., Fruhmann Berger, M., Prakash, D., Csoti, I., Graber, S., Maetzler, W., et al. (2013). Clinical characteristics with an impact on ADL functions of PD patients with cognitive impairment indicative of dementia. *PLoS One* 8:e82902. doi: 10.1371/journal.pone.0082902

Liepelt-Scarfone, I., Graber, S., Kalbe, E., Riedel, O., Ringendahl, H., Schmidt, N., et al. (2021). [Guidelines for the Neuropsychological Assessment of Patients with Parkinson's Disease]. *Fortschr. Neurol. Psychiatr.* 89, 363–373. doi: 10.1055/a-1099-9332

Lin, A., Brook, J., Grill, J. D., and Teng, E. (2017). Participant-Informant Relationships Affect Quality of Life Ratings in Incipient and Clinical Alzheimer Disease. *Am. J. Geriatric. Psychiatry* 25, 297–307. doi: 10.1016/j.jagp.2016.10.007

Litvan, I., Goldman, J. G., Troster, A. I., Schmand, B. A., Weintraub, D., Petersen, R. C., et al. (2012). Diagnostic criteria for mild cognitive impairment in Parkinson's disease: movement Disorder Society Task Force guidelines. *Mov. Disord.* 27, 349–356. doi: 10.1002/mds.24893

Martinez-Martin, P., Benito-Leon, J., Alonso, F., Catalan, M. J., Pondal, M., Tobias, A., et al. (2003). Patients', doctors', and caregivers' assessment of disability using the UPDRS-ADL section: are these ratings interchangeable? *Mov. Disord.* 18, 985–992. doi: 10.1002/mds.10479

Medijainen, K., Paasuke, M., Lukmann, A., and Taba, P. (2015). Functional Performance and Associations between Performance Tests and Neurological Assessment Differ in Men and Women with Parkinson's Disease. *Behav. Neurol.* 2015:519801. doi: 10.1155/2015/519801

Morris, J. C., Mohs, R. C., Rogers, H., Fillenbaum, G., and Heyman, A. (1988). Consortium to establish a registry for Alzheimer's disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. *Psychopharmacol. Bull.* 24, 641–652.

Mosley, P. E., Moodie, R., and Dissanayaka, N. (2017). Caregiver Burden in Parkinson Disease: a Critical Review of Recent Literature. *J. Geriatr. Psychiatry. Neurol.* 30, 235–252. doi: 10.1177/0891988717720302

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x

Pennington, C., Duncan, G., and Ritchie, C. (2021). Altered awareness of cognitive and neuropsychiatric symptoms in Parkinson's disease and Dementia with Lewy Bodies:a systematic review. *Int. J. Geriatr. Psychiatry* 36, 15–30. doi: 10.1002/gps.5415

Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H. Jr., Chance, J. M., and Filos, S. (1982). Measurement of functional activities in older adults in the community. *J. Gerontol.* 37, 323–329. doi: 10.1093/geronj/37.3.323

Pirogovsky, E., Schiehser, D. M., Obtera, K. M., Burke, M. M., Lessig, S. L., Song, D. D., et al. (2014). Instrumental activities of daily living are impaired in Parkinson's disease patients with mild cognitive impairment. *Neuropsychology* 28, 229–237. doi: 10.1037/neu0000045

Pusswald, G., Wiesbauer, P., Pirker, W., Novak, K., Foki, T., and Lehrner, J. (2019). Depression, quality of life, activities of daily living, and subjective memory after deep brain stimulation in Parkinson disease-A reliable change index analysis. *Int. J. Geriatr. Psychiatry* 34, 1698–1705. doi: 10.1002/gps.5184

Ransmayr, G. (2020). [Caregiver burden in patients with Parkinson's disease]. *Fortschr. Neurol. Psychiatr.* 88, 567–572. doi: 10.1055/a-1120-8567

Razani, J., Kakos, B., Orieta-Barbalace, C., Wong, J. T., Casas, R., Lu, P., et al. (2007). Predicting caregiver burden from daily functional abilities of patients with mild dementia. *J. Am. Geriatrics. Soc.* 55, 1415–1420. doi: 10.1111/j.1532-5415.2007.01307.x

Reijnders, J. S., Ehrt, U., Weber, W. E., Aarsland, D., and Leentjens, A. F. (2008). A systematic review of prevalence studies of depression in Parkinson's disease. *Mov. Disord.* 23, 183–189. doi: 10.1002/mds.21803

Sadek, J. R., Stricker, N., Adair, J. C., and Haaland, K. Y. (2011). Performance-based everyday functioning after stroke: relationship with IADL questionnaire and neurocognitive performance. *J. Int. Neuropsychol. Soc.* 17, 832–840. doi: 10.1017/S1355617711000841

Santos Garcia, D., de Deus Fonticoba, T., Cores Bartolome, C., Naya Rios, L., Garcia Roca, L., Martinez Miro, C., et al. (2021). Predictors of Loss of Functional Independence in Parkinson's Disease:results from the COPPADIS Cohort at 2-Year Follow-Up and Comparison with a Control Group. *Diagnostics* 11:1801. doi: 10.3390/diagnostics11101801

Seltzer, B., Vasterling, J. J., Mathias, C. W., and Brennan, A. (2001). Clinical and neuropsychological correlates of impaired awareness of deficits in Alzheimer disease and Parkinson disease: a comparative study. *Neuropsychiatry Neuropsychol. Behav. Neurol.* 14, 122–129.

Shulman, L. M., Pretzer-Aboff, I., Anderson, K. E., Stevenson, R., Vaughan, C. G., Gruber-Baldini, A. L., et al. (2006). Subjective report versus objective measurement of activities of daily living in Parkinson's disease. *Mov. Disord.* 21, 794–799. doi: 10.1002/mds.20803

Skinner, J. W., Lee, H. K., Roemmich, R. T., Amano, S., and Hass, C. J. (2015). Execution of Activities of Daily Living in Persons with Parkinson Disease. *Med. Sci. Sports Exerc.* 47, 1906–1912. doi: 10.1249/MSS.0000000000000598

Sturm, W., Willmes, K., and Horn, W. (1993). *Leistungsprüfsystem für 50-90jährige (LPS 50+)*. Göttingen: Hogrefe.

Sulzer, P., Baumer, A., Hoang, H. G., Becker, S., Lonneker, H. D., Graessel, E., et al. (2020). Assessment of Cognitive-Driven Performance-Based Activities of Daily Living Dysfunction in Parkinson's Disease. *J. Int. Neuropsychol. Soc.* 26, 430–440. doi: 10.1017/S1355617719001322

Tomlinson, C. L., Stowe, R., Patel, S., Rick, C., Gray, R., and Clarke, C. E. (2010). Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov. Disord.* 25, 2649–2653. doi: 10.1002/mds.23429

Yamawaki, R., Nankaku, M., Kusano, Y., Tajima, A., Ikeguchi, R., and Matsuda, S. (2020). Evaluation of Cognitive Function in Relation to Progression of Parkinson Disease. *Am. J. Phys. Med. Rehabil.* 99, 626–629. doi: 10.1097/PHM.0000000000001385

Zanetti, O., Geroldi, C., Frisoni, G. B., Bianchetti, A., and Trabucchi, M. (1999). Contrasting results between caregiver's report and direct assessment of activities of daily living in patients affected by mild and very mild dementia:the contribution of the caregiver's personal characteristics. *J. Am. Geriatrics Soc.* 47, 196–202. doi: 10.1111/j.1532-5415.1999.tb04578.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 22 March 2022
doi: 10.3389/fnagi.2022.808518

Investigating the Impact of Environment and Data Aggregation by Walking Bout Duration on Parkinson's Disease Classification Using Machine Learning

*Rana Zia Ur Rehman[1], Yu Guan[2], Jian Qing Shi[3,4], Lisa Alcock[1], Alison J. Yarnall[1,5], Lynn Rochester[1,5] and Silvia Del Din[1]**

[1] Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, United Kingdom, [2] School of Computing, Newcastle University, Newcastle upon Tyne, United Kingdom, [3] School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne, United Kingdom, [4] Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen, China, [5] The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, United Kingdom

OPEN ACCESS

Edited by:
Ondrej Bezdicek, Charles University, Czechia

Reviewed by:
Federico Villagra, Aberystwyth University, United Kingdom
Thomas Schauer, Technical University of Berlin, Germany

***Correspondence:**
Silvia Del Din
silvia.del-din@ncl.ac.uk

Specialty section:
This article was submitted to Parkinson's Disease and Aging-related Movement Disorders, a section of the journal Frontiers in Aging Neuroscience

Received: *03 November 2021*
Accepted: *14 February 2022*
Published: *22 March 2022*

Citation:
Rehman RZU, Guan Y, Shi JQ, Alcock L, Yarnall AJ, Rochester L and Del Din S (2022) Investigating the Impact of Environment and Data Aggregation by Walking Bout Duration on Parkinson's Disease Classification Using Machine Learning. Front. Aging Neurosci. 14:808518. doi: 10.3389/fnagi.2022.808518

Parkinson's disease (PD) is a common neurodegenerative disease. PD misdiagnosis can occur in early stages. Gait impairment in PD is typical and is linked with an increased fall risk and poorer quality of life. Applying machine learning (ML) models to real-world gait has the potential to be more sensitive to classify PD compared to laboratory data. Real-world gait yields multiple walking bouts (WBs), and selecting the optimal method to aggregate the data (e.g., different WB durations) is essential as this may influence classification performance. The objective of this study was to investigate the impact of environment (laboratory vs. real world) and data aggregation on ML performance for optimizing sensitivity of PD classification. Gait assessment was performed on 47 people with PD (age: 68 ± 9 years) and 52 controls [Healthy controls (HCs), age: 70 ± 7 years]. In the laboratory, participants walked at their normal pace for 2 min, while in the real world, participants were assessed over 7 days. In both environments, 14 gait characteristics were evaluated from one tri-axial accelerometer attached to the lower back. The ability of individual gait characteristics to differentiate PD from HC was evaluated using the Area Under the Curve (AUC). ML models (i.e., support vector machine, random forest, and ensemble models) applied to real-world gait showed better classification performance compared to laboratory data. Real-world gait characteristics aggregated over longer WBs (WB 30–60 s, WB > 60 s, WB > 120 s) resulted in superior discriminative performance (PD vs. HC) compared to laboratory gait characteristics ($0.51 \leq AUC \leq 0.77$). Real-world gait speed showed the highest AUC of 0.77. Overall, random forest trained on 14 gait characteristics aggregated over WBs > 60 s gave better performance (F1 score = 77.20 ± 5.51%) as compared to laboratory results (F1 Score = 68.75 ± 12.80%). Findings from this study suggest that the choice of

environment and data aggregation are important to achieve maximum discrimination performance and have direct impact on ML performance for PD classification. This study highlights the importance of a harmonized approach to data analysis in order to drive future implementation and clinical use.

Clinical Trial Registration: [09/H0906/82].

Keywords: Parkinson's disease, gait, real-world, accelerometer, machine learning, laboratory, gait aggregation, wearables

INTRODUCTION

Parkinson's disease (PD) is the second most common neurodegenerative disease after Alzheimer's disease (Nutt and Wooten, 2005; Feigin et al., 2019). PD prevalence has doubled over the past 25 years and now affects approximately 10 million people worldwide (Dorsey et al., 2018). Due to the progressive nature of PD (Dorsey et al., 2013; Emamzadeh and Surguchov, 2018), both motor (Hobert et al., 2019) and non-motor (Przedborski et al., 2003; Jankovic, 2008) symptoms have a significant impact on quality of life and increased burden on healthcare costs (von Campenhausen et al., 2011). Currently, diagnostic criteria for PD are based on motor features assessed with clinical scales (Jankovic, 2008; Postuma et al., 2015). However, the diagnostic accuracy of PD in a clinical setting is only 74% if performed by a non-expert and 80% by a movement disorder specialist (Rizzo et al., 2016). Given the relatively low rates of accurate diagnosis, particularly in the early stages, there is a need for additional diagnostic aids (Mancini et al., 2011). The application of gait analysis may be a promising addition to the diagnostic toolkit (Buckley et al., 2019; Viceconti et al., 2020).

Previous work has shown that an objective gait assessment obtained in lab settings and the clinic can be used to classify PD using machine learning (ML) approaches (Rehman et al., 2019a,b, 2020a,b). However, assessing gait in both the lab and the clinic has some key limitations. The patient is required to attend specialist facilities, and assessments often do not represent the range of challenges associated with habitual walking (Orendurff et al., 2008). Moreover, individuals tend to perform better (walk faster) during performance tests which reflects walking capacity ("can do") (Del Din et al., 2016a) compared to everyday life which captures the functional performance ("actually do") of the participant (Hillel et al., 2019; Maetzler et al., 2020; Shah et al., 2020c; Warmerdam et al., 2020; Atrsaei et al., 2021). The real world may therefore provide a more sensitive and pragmatic context to identify and classify PD (Shah et al., 2020c). Increasing interest in the use of inertial measurement units (IMUs) to monitor gait in people with PD in the lab is evident (González et al., 2010; McCamley et al., 2012; Zijlstra and Zijlstra, 2013; Godfrey et al., 2015; Del Din et al., 2016c), as is monitoring gait continuously in the real world over multiple consecutive days (De Bruin et al., 2007; Weiss et al., 2013, 2014; Godfrey et al., 2014). However, several methodological challenges remain for a better understanding and analysis of real-world gait data. These include extraction of relevant gait characteristics and appropriate use of data aggregation for analysis, e.g., averaging gait characteristics using various WB durations.

Spatiotemporal gait characteristics [from the gait domains of pace, rhythm, variability, asymmetry, and postural control (Lord et al., 2013)] from lab and real-world data are significantly different in people with PD compared to healthy controls (HCs) (Maetzler et al., 2013; Del Din et al., 2016a). However, methods for analysis of data obtained in real-world settings rely on selecting the protocol for gait assessment (e.g., environment and duration) and data aggregation by walking bout (WB) duration (e.g., aggregating all WB's or selecting an optimal bout duration) (Del Din et al., 2016a,b, 2019; Shah et al., 2020a; Warmerdam et al., 2020). All these options impact on the quantification of spatiotemporal gait characteristics and subsequent results (Del Din et al., 2016b).

Real-world gait consists of a variety of WBs of different durations, the majority of which are short (<10 s, approximately 50%) with only 3% over 60 s for both PD and HC (Del Din et al., 2016a). In contrast, lab-based gait assessments are based on standardized walking distances such as 4 or 10 m (Del Din et al., 2016a,c; Van Ancum et al., 2019) or duration (e.g., 2 min) (Rehman et al., 2019a,b; Del Din et al., 2020). Comparison of data obtained in the lab and in the real-world is therefore challenging.

In previous work, ML classifiers have been trained on data from lab-based gait assessments (Rehman et al., 2019a,b, 2020a,b). The impact of environment (lab vs. real-world) and data aggregation by WB duration on PD classification has not been thoroughly explored. Different WB durations also influence the distribution of gait characteristics. Therefore, ML models need to be able to account for multiple distributions (due to inclusion of a variety of short and long WBs) of real-world gait characteristics. To the best of the authors' knowledge, the impact of WB durations on the classification of PD using machine learning approaches has not yet been investigated.

The aims of this study are therefore to: (i) investigate the impact of environment (gait assessment in lab vs. real world) and (ii) data aggregation by WB duration on gait characteristics and performance of ML models to accurately classify PD. Based on current available univariate analyses (Del Din et al., 2016a, 2019; Shah et al., 2020c), we hypothesized that: (i) real-world gait would be more sensitive for performing the ML based classification of PD compared to lab gait assessment; (ii) associations between lab-based and real-world gait would vary depending on WB duration; and (iii) ML model performance would be influenced by WB duration.

MATERIALS AND METHODS

Participants

In this cross-sectional analysis, 52 HCs and 47 people with PD were included from the 18 month time point of the "Incidence of Cognitive Impairment in Cohorts with Longitudinal Evaluation–GAIT" (ICICLE-GAIT) study (Khoo et al., 2013; Yarnall et al., 2014). ICICLE-GAIT is a study nested within ICICLE-PD which recruited participants between June 2009 and December 2011 (Khoo et al., 2013). PD participants were recruited from local movement disorders clinics (Khoo et al., 2013) and had a diagnosis of idiopathic PD according to the United Kingdom Brain Bank Criteria (Khoo et al., 2013; Yarnall et al., 2014). PD participants who have Parkinsonism disorders or an atypical form of Parkinson's disease, with poor knowledge of working English language, or with cognitive impairment (Mini-Mental State Examination score < 24) were excluded from the study. The HC participants were recruited from the local community and included provided that they were able to walk independently and were without significant motor, mood, or cognitive impairment. ICICLE-GAIT received ethical approval from the Newcastle and North Tyneside Research Ethics Committee (REC No. 09/H0906/82). Study procedures were conducted according to the Declaration of Helsinki and all participants gave written informed consent prior to participating.

Demographics and Clinical Characteristics

Participant demographics such as sex, age, mass, height, and BMI were recorded. The Montreal Cognitive Assessment (MoCA) was used to measure global cognition (Nasreddine et al., 2005). Balance confidence was assessed using the Activities Specific Balance Confidence scale (ABC) (Powell and Myers, 1995). PD motor severity was assessed with Part III of the Movement Disorder Society Unified Parkinson's Disease Rating Scale (MDS-UPDRS III) (Goetz et al., 2008) and disease stages were also recorded according to Hoehn and Yahr (1998). The levodopa equivalent daily dose (LEDD; mg/day) was also calculated (Tomlinson et al., 2010; Lawson et al., 2016).

Gait Assessment

Equipment and Protocol

Gait assessment was performed in the lab and real-world using a tri-axial (x: vertical, y: mediolateral, z: anteroposterior) accelerometer (Axivity, AX3, sample frequency: 100 Hz, range: ± 8 g) on the lower back, as shown in **Figure 1A**. In the lab, a 2 min continuous walk around an oval circuit was performed (Rehman et al., 2019b). PD participants' gait was assessed while optimally medicated (approximately one hour after medication intake). In the real-world, gait was monitored continuously for 7 days (Del Din et al., 2016a,b, 2019). This took place following the lab assessment. Participants were instructed to perform their usual activities. Further details can be found in previous work (Godfrey et al., 2014; Del Din et al., 2016a, 2019).

Data Processing and Extraction of Gait Characteristics–Lab

Data from the accelerometer was downloaded to a computer for offline processing in MATLAB (R2019a). The vertical component of the transformed acceleration signal was filtered first to 20 Hz with a 4th order Butterworth filter (Moe-Nilssen, 1998; Zijlstra and Hof, 2003; McCamley et al., 2012). To detect the events within each gait cycle (**Figure 1B**), the initial contact (IC, heel strike) and final contact (FC, toe-off) points were identified with the help of a Gaussian continuous wavelet transform. Additional temporal gait characteristics (step time, swing time, and stance time) were quantified based on IC and FC (McCamley et al., 2012; Godfrey et al., 2015; Del Din et al., 2016c). For the evaluation of spatial characteristics (step length) the inverted pendulum model was utilized (Zijlstra and Zijlstra, 2013) and step velocity was calculated as the ratio of step length and step time (Del Din et al., 2016c). Variability was calculated as the standard deviation from all steps and asymmetry as the absolute difference of alternative steps (left and right) (Del Din et al., 2016c). The detailed method for the evaluation of spatiotemporal gait characteristics is described in previous work (Lord et al., 2013; Godfrey et al., 2015; Del Din et al., 2016c).

Fourteen spatiotemporal gait characteristics (**Figure 1C**) were extracted based on ICs and FCs and mapped onto five domains: pace (step velocity, step length, swing time variability), rhythm (step time, swing time, stance time), variability (step velocity variability, step length variability, step time variability, swing time variability, stance time variability), asymmetry (step time asymmetry, swing time asymmetry, stance time asymmetry), and postural control (step length asymmetry) (Lord et al., 2013; Godfrey et al., 2015; Del Din et al., 2016c).

Data Processing, Extraction of Gait Characteristics, and Data Aggregation–Real-World

For the real-world gait assessment, data was downloaded to a computer for offline processing in MATLAB (R2019a). Accelerometry data was segmented into each calendar day and WBs were detected based on the magnitude and standard deviation of the acceleration signal (Del Din et al., 2016d; Hickey et al., 2016). A WB was defined as the continuous length of time spent during walking (Godfrey et al., 2014), with at least three steps (Del Din et al., 2016a, 2019). No resting period thresholds between consecutive WBs were set so that each WB was individually considered (and not merged to other WBs) (Barry et al., 2015). Gait characteristics were firstly evaluated for each WB by combining all steps within a WB. Then, all WBs were combined for each day to provide a daily average. Finally, each day was combined to provide a 7 day average for each gait characteristic (Del Din et al., 2016a, 2019). The same fourteen gait characteristics were extracted from the real-world (Lord et al., 2013; Godfrey et al., 2015; Del Din et al., 2016c) for comparison with lab-based gait (**Figure 1C**).

To investigate the impact of real-world data aggregation by WB duration on gait characteristics and ML models, a comprehensive approach was adopted. WB of various durations (seconds) were considered and aggregated over the 7 days (**Figure 1A**). In total, fourteen WB durations were chosen,

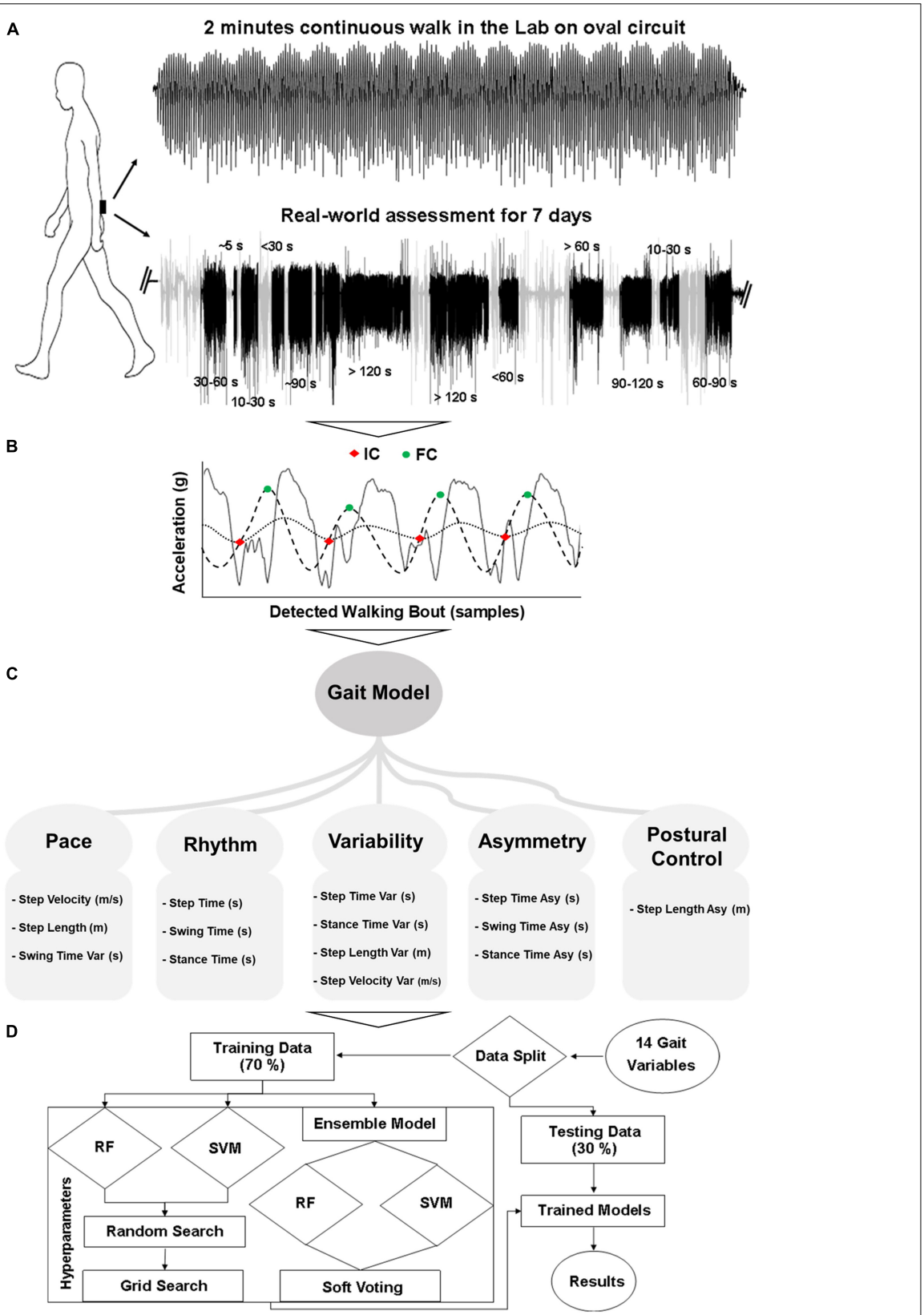


FIGURE 1 | Overall workflow from gait assessment to classification: **(A)** Gait assessment protocol, **(B)** WB detection and gait characterization, **(C)** 14 gait characteristics (Var, variability; Asy, asymmetry), **(D)** Classification modeling.

and the average of all WBs was used to describe each gait characteristic. The six most optimal and distinct WB durations without having redundant information by combining the incremental WBs (i.e., WBs $\leq$ 10 s, 10 $<$ WBs $\leq$ 30 s, 30 $<$ WBs $\leq$ 60 s, 60 $<$ WBs $\leq$ 120 s, WBs $>$ 60 s, WBs $>$ 120 s) are presented in the manuscript to reduce the data for clear message. However, the remaining WB durations (i.e., WBs $\leq$ 5 s, WBs $\leq$ 30 s, WBs $\leq$ 60 s, WBs $\leq$ 90 s, WBs $\leq$ 120 s, 5 $<$ WBs $\leq$ 10 s, 60 $<$ WBs $\leq$ 90 s, 90 $<$ WBs $\leq$ 120 s) are provided in the **Supplementary Material**.

Statistical Analysis

Normality of the data (gait characteristics) was checked by plotting histograms and using the Shapiro Wilk test for each environmental condition. In addition, rain clouds and box plots were used to visually check the distribution for each group and shift in distribution among groups (PD vs. HC) for each gait characteristic. To evaluate differences between PD, HC, and the impact of environment, a mixed ANOVA was performed on the data aggregation by WB duration and their combined effect (interaction) on each gait characteristic. Based on the data distribution, student *t*-test and Mann Whitney *U*-test were used to evaluate differences between the PD and HC groups. Given the exploratory nature of this analysis, we used a p-value < 0.05 to guide statistical interpretation and did not make adjustments for multiple comparisons (Rothman, 1990; Perneger, 1998). This is due to the inclusion of mixed ANOVA for overall statistics and area under the receiver operating characteristics curve (AUC) analysis for each gait characteristic to investigate its discriminatory power (PD vs. HC) under different environmental conditions and aggregation by WB duration. In addition, the p-values are provided to assess the statistical significance of between group differences. The relationship between lab and real-world gait characteristics was also assessed with the Pearson's correlation coefficient.

Machine Learning Classification of Parkinson's Disease

Three different ML models were used: support vector machine (SVM), random forest (RF), and an ensemble of these two classifiers (**Figure 1D**; Rehman et al., 2019a,b). The ensemble model made the decision based on soft voting (probability) (Pedregosa et al., 2011). Due to the variety of data distributions, instead of training a separate model for each WB threshold, ML models were trained by combining all WB duration data. This allowed one single model to be developed, which could cater for the distribution of the entire dataset. Training performance of the models was evaluated using a 10-fold cross-validation technique on 70% of data, and separate testing was done on each WB duration threshold by keeping 30% of the data for testing. This rigorous training and testing process was repeated 10 times based on different random seed values. Classifier performance was evaluated in terms of accuracy, F1 score, AUC, sensitivity, and specificity (Rehman et al., 2019a, 2020a,b). In addition, influential gait characteristics were also identified based on their importance in RF and recursive feature elimination (RFE) technique with SVM-linear (Rehman et al., 2019a,b). Model hyperparameters were optimized with grid search. The standard python library SciKit learn was used for ML analysis.

RESULTS

Demographic and clinical characteristics are summarized in **Table 1**. There were no significant differences between the PD and HC groups for sex, age, height, mass, and BMI. People with PD had lower cognitive scores (MoCA) and reduced Activities-Specific Balance Confidence (ABC) score compared to HC ($p < 0.05$). PD participants had an average disease duration of 26 months, the majority of which were Hoehn and Yahr stage II with an average MDS-UPDRS III score of 31.5 $\pm$ 9.8.

Impact of Environment and Data Aggregation on Gait Characteristics

Overall Statistics

The distribution of WB depending on duration is shown in **Figure 2** and the distribution of step velocity is shown in **Figure 3**. Distributions for the remainder of gait characteristics are reported in **Supplementary Figures 1, 2**. Overall, mixed ANOVA statistical analysis results are presented in **Table 2**. Statistical differences between PD and HC are displayed with a heat map in **Figure 4**.

There was a significant interaction (**Table 2**) between environment (lab and real-world, including data aggregation by

TABLE 1 | Demographics and clinical measures of the Parkinson's disease (PD) and healthy controls (HC) group.

Characteristics	HC ($n = 52$) Mean ± SD	PD ($n = 47$) Mean ± SD	p
M/F (*n*)	28/24	32/15	0.083
Age (years)	70.39 ± 6.88	68.36 ± 8.98	0.216
Height (m)	1.69 ± 0.08	1.70 ± 0.08	0.542
Mass (kg)	81.13 ± 15.15	80.27 ± 15.67	0.786
BMI (kg/m^2)	28.29 ± 4.23	27.62 ± 4.62	0.455
MoCA	27.61 ± 2.39	26.28 ± 3.60	**0.037**
ABCs (0–100%)	91.02 ± 11.65	80.88 ± 16.18	**0.003**
Medication (LEDD, mg/day)		415.08 ± 212.61	
Time from Clinical Diagnosis (months)		26.42 ± 5.48	
Hoehn and Yahr (*n*)		HY I–7 (15%)	
		HY II–38 (81%)	
		HY III–2 (4%)	
MDS-UPDRS III		31.53 ± 9.79	
		(HY I–16.6 ± 4.73)	
		(HY II–33.28 ± 8.81)	
		(HY III–35.5 ± 0.71)	

M, Males; F, Females; BMI, Body Mass Index; MoCA, Montreal Cognitive Assessment; ABCs, Activities Specific Balance Confidence scale; LEDD, Levodopa Equivalent Daily Dose; MDS-UPDRS, Movement Disorder Society Unified Parkinson's Disease Rating Scale. Bold values mean a significant difference between PD and HC.

TABLE 2 | Mixed ANOVA results: main effects and interactions between gait assessment environmental conditions (lab vs. real-world) and groups (PD vs. HC) for each gait characteristic.

Gait characteristics	Between participant factor: Group (HC, PD)			Within participant factor: Gait data (Lab, real-world, and all WB)			Interaction:Group × Gait		
	F	*p*	η^2	*F*	*p*	η^2	*F*	*p*	η^2
Step Velocity (m/s)	9.46	0.003*	0.089	73.36	<0.001*	0.431	5.47	0.001*	0.053
Step Length (m)	9.95	0.002*	0.093	138.22	<0.001*	0.588	2.56	0.058	0.026
Swing Time Variability (s)	6.83	0.01*	0.066	158.35	<0.001*	0.62	2.25	0.085	0.23
Step Time (s)	5.45	0.022*	0.053	42.58	<0.001*	0.305	4.83	0.001*	0.047
Swing Time (s)	6.85	0.01*	0.066	66.63	<0.001*	0.407	3.67	0.007*	0.036
Stance Time (s)	4.64	0.034*	0.046	37.07	<0.001*	0.276	4.54	0.002*	0.045
Step Velocity Variability (m/s)	0.22	0.638	0.002	128.51	<0.001*	0.57	1.49	0.218	0.015
Step Length Variability (m)	1.58	0.211	0.016	58.31	<0.001*	0.375	2.11	0.097	0.021
Step Time Variability (s)	5.15	0.026*	0.5	173.13	<0.001*	0.641	2.81	0.04*	0.028
Stance Time Variability (s)	5.79	0.018*	0.056	170.86	<0.001*	0.638	3.14	0.023*	0.031
Step Time Asymmetry (s)	4.44	0.038*	0.044	1018.75	<0.001*	0.913	1.11	0.342	0.011
Swing Time Asymmetry (s)	7.79	0.006*	0.074	1117.42	<0.001*	0.92	2.64	0.047*	0.026
Stance Time Asymmetry (s)	5.05	0.027*	0.05	986.81	<0.001*	0.911	2.04	0.113	0.021
Step Length Asymmetry (m)	0.18	0.671	0.002	911.29	<0.001*	0.904	2.29	0.051	0.023

*Partial eta squared (η^2) represents effect size; F statistic (F); significance (p) indicated as *.*

WB duration) and group (HC, PD) for seven characteristics (step velocity, step time, swing time, stance time, step time variability, stance time variability, and swing time asymmetry). However, the main effect of within participant factor revealed that all the gait characteristics evaluated in the lab setting were significantly different from those evaluated in the real-world setting and for all data aggregations. Similarly, the main effect of between-participant factor revealed that there were significant differences between PD and HC for all the gait characteristics except step velocity variability, step length variability, and step length asymmetry.

Impact of Environment [Lab vs. Real-World (All Walking Bouts)] on Between-Group Differences (Parkinson's Disease vs. Healthy Control)

Lab vs. Real-World (All Walking Bouts)

During the 2-min continuous walk in the lab, both groups walked faster with a shorter step time and longer step length compared to when in the real-world regardless of WB duration. Even for the longest WB (>120 s), these characteristics (step velocity and step length) were reduced compared to the 2-min (120 s) continuous walk in the lab. As an example, the distribution of step velocity for both groups and environmental conditions under different data aggregation by WB durations is shown in **Figure 3**. Gait variability (swing time, step time, stance time, step length, and step velocity) was reduced when measured in the lab compared to the real-world. Gait was more symmetrical when measured in the lab (step time, stance time, swing time, and step length) for both PD and HC compared to real-world gait irrespective of WB duration.

Parkinson's Disease vs. Healthy Control

Both in the lab and real-world conditions, when combining all WB durations, PD participants walked slower, with slower step time and shorter step length compared to HC (**Supplementary Figure 1**). Step velocity and step length were significantly different ($p < 0.05$) between PD and HC in both lab and real-world conditions (**Figure 4**). In the real-world, PD participants had significantly lower swing time variability compared to HC ($p = 0.033$). None of the asymmetry-based gait characteristics measured in the lab were significantly different between PD and HC, except for step length asymmetry (**Figure 4**). Similarly, for asymmetry-based characteristics in the real-world, only swing time asymmetry, based on the combination of all WBs, was significantly higher for PD compared to HC ($p = 0.006$) (**Figure 4**).

Effect of Real-World Data Aggregation on Gait

Distribution of Data Aggregation by Walking Bout Durations

All PD and HC participants had WBs across all duration thresholds. The distribution of WBs are shown in **Figure 2** and **Supplementary Figure 3**. The majority of WBs (87% for PD and 85% for HC) were of shorter duration (≤10 s), with relatively few WB per day found over 120 s (1.8% for PD and 1.9% for HC). Overall, HC had a greater number of WB (total $n = 170{,}368$ from 52 HC) over 7 days of continuous assessment compared to the PD group (total $n = 134{,}156$ from 47 PD).

Effect of Data Aggregation by Walking Bout Durations on Between Group Differences

For both PD and HC groups, the slowest speed was observed during very short WBs (≤10 s) compared to long WBs > 10 s (**Figure 3** and **Supplementary Figures 2, 4**). The most significant ($p < 0.01$) group differences between PD and HC were found in longer WBs (>60 s or 120 s) as compared to shorter WBs (**Figure 4**). Similarly, reduced step length and shorter step time were observed in short WBs (≤10 s) as compared to long WBs. Step time was significantly slower for PD in longer WBs such

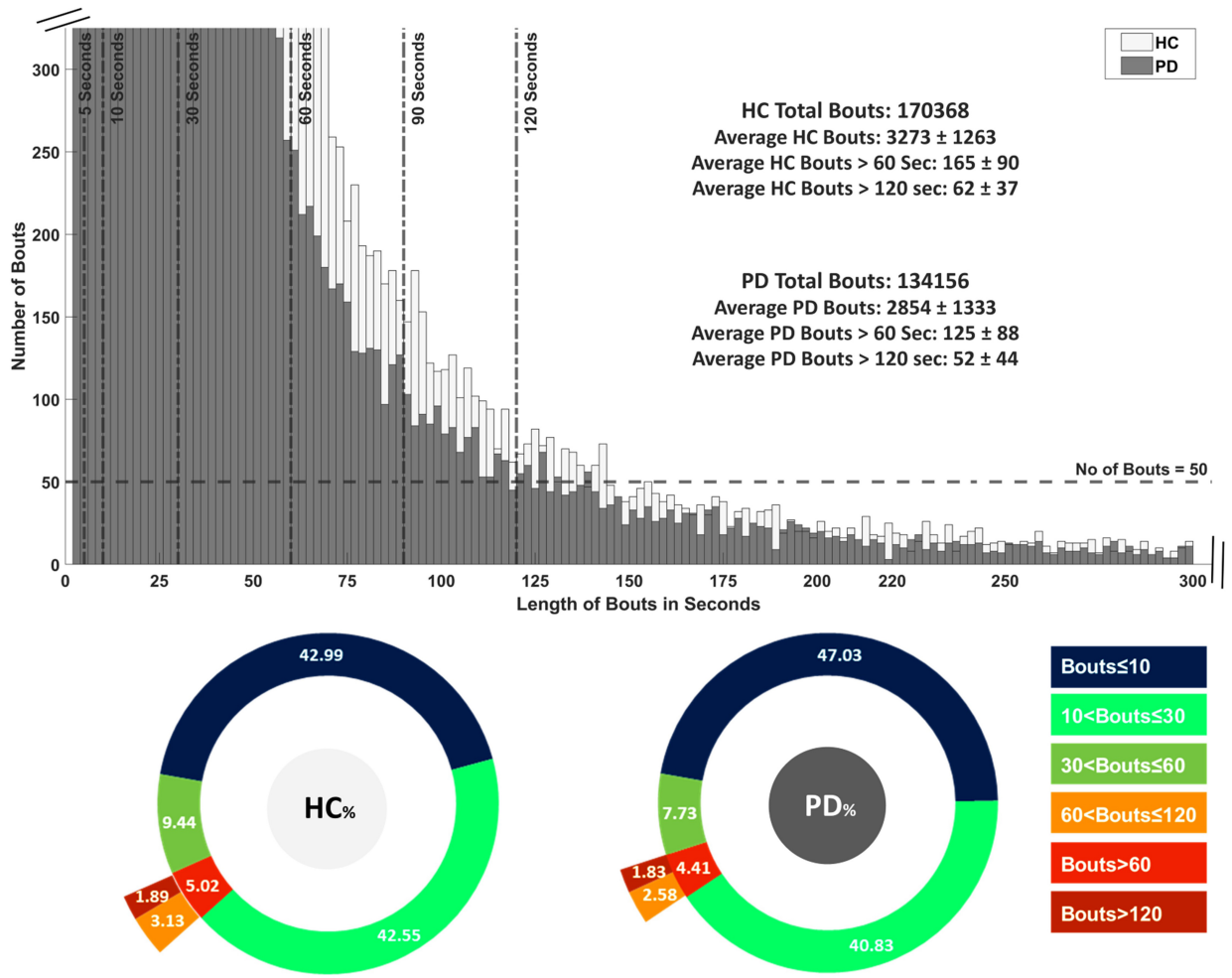


FIGURE 2 | Distribution of all detected walking bouts (WBs) in Parkinson's disease (PD) and healthy control (HC) groups in the real-world assessment of 7 days. WB are categorized intro 14 thresholds based on time in seconds followed by their average of 7 days.

as >60 s ($p = 0.001$) and >120 s ($p < 0.001$) compared to HC. Interestingly, gait variability in the longer WBs (>60 s and >120 s) resulted in significant differences between the groups for swing time variability, step length variability, step time variability, and stance time variability. All the asymmetry-based gait characteristics behaved differently in short and long WBs. For example, in the longer WBs (>120 s), asymmetry-based gait characteristics were similar or close to lab-based gait asymmetry characteristics.

To summarize, in the real world, significant group differences (PD vs. HC) were identified for WBs longer than 60 or 120 s for all gait characteristics apart from asymmetry-based gait characteristics.

Association Between Gait Characteristics Collected in the Lab vs. Real-World

The association between lab and real-world gait characteristics is shown in **Figure 5**. In general, lab and real-world gait characteristics showed either no correlation or weak-to-moderate association with one another. However, stronger correlations were noted for the PD group compared to HC with correlations >0.5 observed for step length, step time, and stance time. Stance time resulted in the strongest correlations, followed by step time, compared to other gait characteristics, with the strongest correlation of 0.59 observed for WBs of 10–30 s and 30–60 s. Real-world gait speed resulted in a weak correlation for both PD and HC (max 0.388) with lab-based gait speed in the longer WBs > 60 or >120 s. Short WBs (≤10 s) had weak correlation compared to longer WBs (>10 s). Variability characteristics were negatively correlated between the lab and real-world gait assessments. Results for all WB durations are presented in **Supplementary Figure 5**.

Impact of Environment and Data Aggregation on Classification of Parkinson's Disease

Results showing the impact of environment and data aggregation by WB durations on each individual gait characteristic

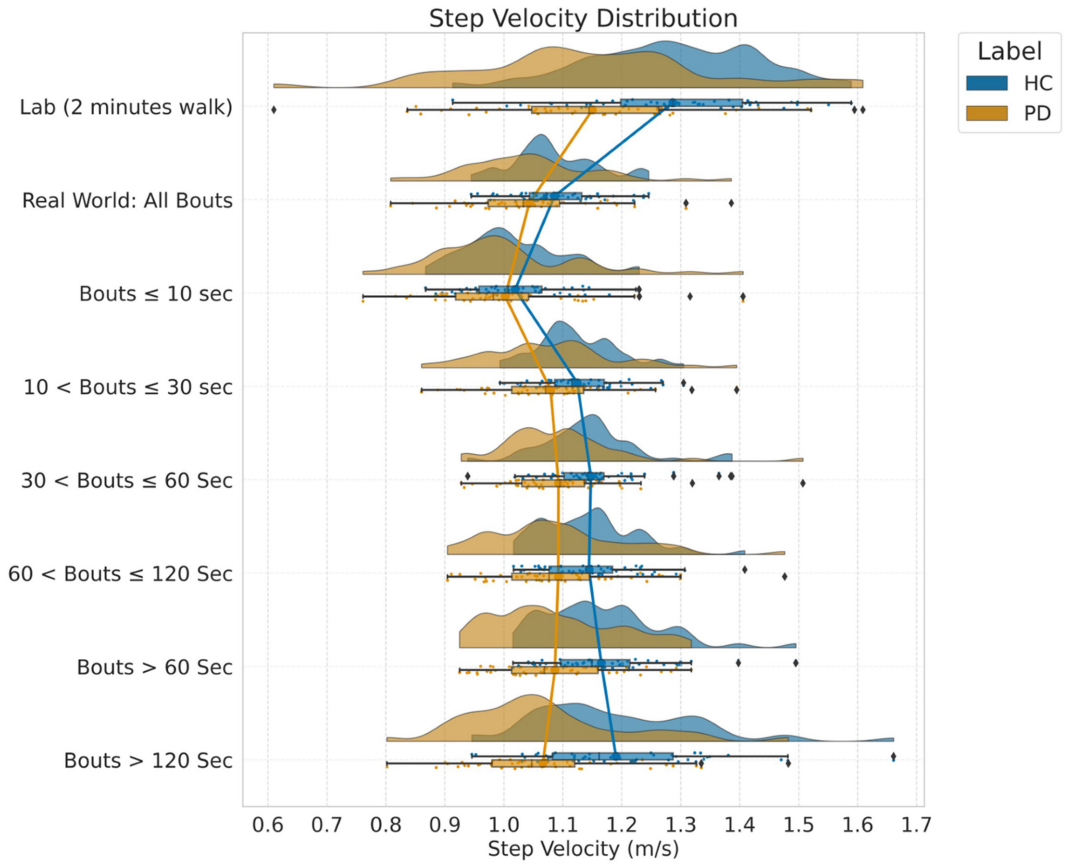


FIGURE 3 | Distribution of step velocity (m/s) into different WB thresholds and average of 7 days.

and ML models for classification of PD are presented in **Figures 6**, **7** (**Supplementary Figures 6–8**) and **Table 3** (**Supplementary Table 1**).

Machine Learning Performance in the Lab vs. Real-World

Based on the AUC for each gait characteristic discriminating PD from HC (**Figure 6**), real-world gait characteristics combining all WBs had relatively low AUC values compared to gait assessed in lab settings. However, the asymmetry-based characteristics in lab had lower AUC (0.51–0.57) as compared to real-world asymmetry-based gait characteristics (AUC = 0.61–0.68).

For the classification of PD, various gait characteristics were statistically significant between the groups (PD vs. HC) in lab and real-world settings. Therefore, the classifiers were trained on the overall 14 gait characteristics. During the training phase, performance of the classifiers was based on the 10-fold cross-validation and ranged between 72 and 95% (**Supplementary Figure 7**). These trained classifiers were tested separately on the 30% average lab and real-world test data. Random forest performed better under both environment conditions. When combining all WBs, real-world gait gave lower classification performance (accuracy: 60.67 ± 10.65) compared to lab-based gait (accuracy: 64.67 ± 15.02). This lower performance from real-world data (all WBs combined) was observed (**Figure 7** for F1 score) in all the three models (random forest, support vector machine, and ensemble model). However, only RF performance was statistically significant from SVM under both environmental conditions.

Impact of Data Aggregation by Walking Bout Duration on Machine Learning Performance

Real-world gait characteristics had higher AUC compared to lab-based gait assessment for selected WB durations (**Figure 6**). The maximum AUC of 0.765 was observed for step velocity in real-world gait assessment from longer WBs (>120 s), followed by the lab-based step velocity (AUC of 0.721), 30 < WBs ≤ 60 s (AUC of 0.714), and WBs > 60 s (AUC of 0.709). All the rhythm-based gait characteristics (step time, swing time, and stance time) had an AUC around 0.7 when aggregated across longer WBs > 60 s or > 120 s. A maximum AUC of 0.703 was found in 30 < WBs ≤ 60 s for swing time asymmetry. To summarize, the maximum AUC from real-world gait characteristics were found in the longer WBs (30 < WBs ≤ 60 s, WBs > 60 s, and WBs > 120 s).

Data distribution for different thresholds of WB duration varies. Therefore, classifiers were trained on the combined data to accommodate all distributions. In addition, because

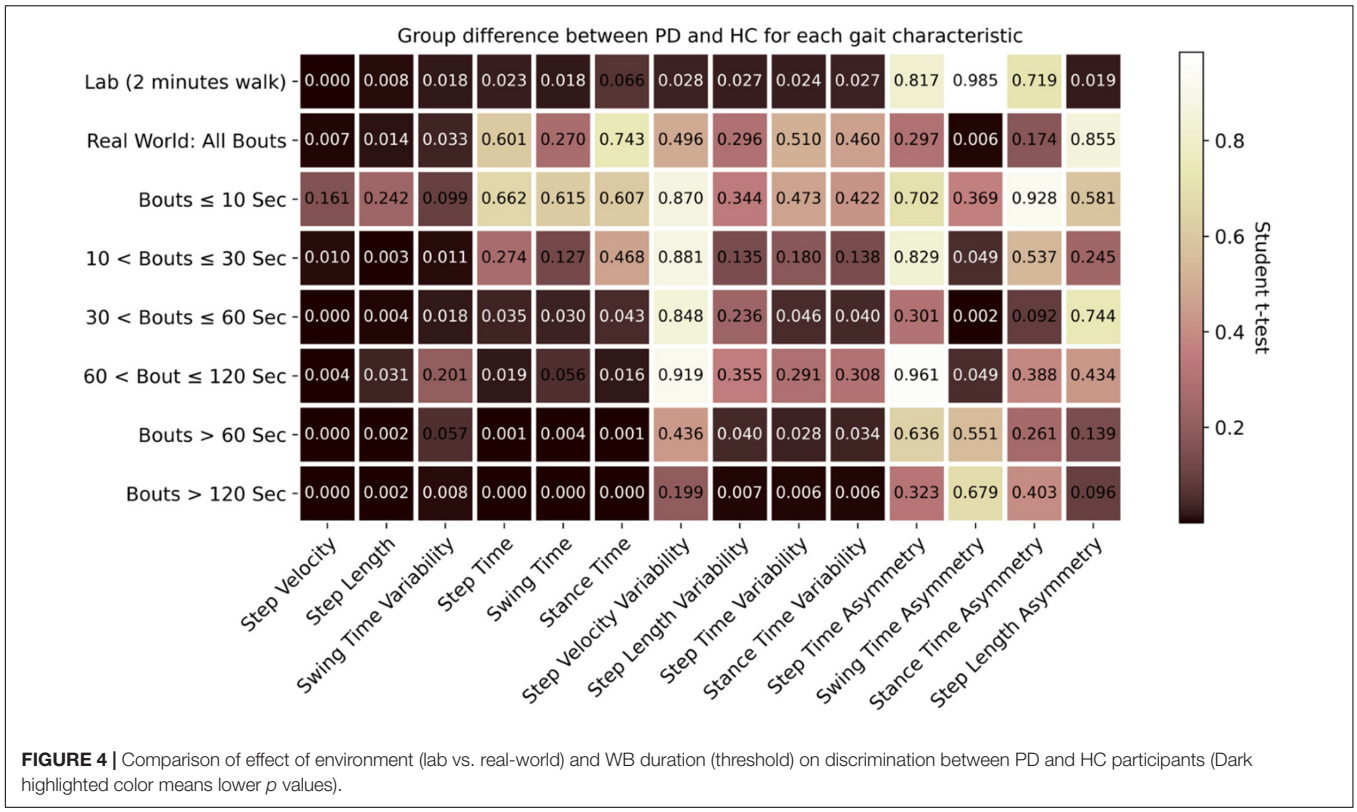


FIGURE 4 | Comparison of effect of environment (lab vs. real-world) and WB duration (threshold) on discrimination between PD and HC participants (Dark highlighted color means lower *p* values).

there were differences in gait characteristics (discriminating PD vs. HC) for different thresholds of WB duration, classifiers were trained on all 14 gait characteristics. During the training phase, the performance of the classifier based on the 10-fold cross-validation ranged between 72 and 95% (**Supplementary Figure 7**). Overall, RF performed better on the new data set (30%) used for testing as compared to other classifiers (**Figure 7**). In addition, RF classification performance was significantly different from SVM in all WB durations except WBs < 10 s, while RF was only significantly different from ensemble model in $10 <$ WBs < 30 s and WBs > 120 s. Longer WBs > 60 s gave a better classification performance in discriminating PD from HC as compared to other WB durations in the real-world data. However, the classifier performance varies with different data aggregation by WB durations, which indicates that comparing gait performance from the same participants in different environments (and WB durations) can influence the classifiers. Maximum testing performance of the classifiers were obtained from $30 <$ WBs ≤ 60 s, WBs > 60 s, and WBs > 120 s (**Table 3**).

Influential gait characteristics were similar to those characteristics with a higher AUC (e.g., step velocity, step length, step time, swing time, stance time, and swing time variability; **Supplementary Figure 9**). Based on the importance of characteristics in the RF classifier, swing time, step velocity, stance time, swing time variability, and step length were identified as the top five characteristics. Similarly, based on the RFE with support vector machine, step velocity, step length, stance time, step time, and swing time were identified. Based on the common characteristics in the top five, step velocity, step length, swing time, and stance time were identified by both classifiers.

DISCUSSION

This is the first study to comprehensively investigate the impact of environment and data aggregation by WB duration on ML performance for the classification of PD. Based on the results, environment and aggregation of real-world data by WB duration influenced each individual gait characteristics for both groups and subsequent performance of ML models. We found a weak to moderate association between lab and real-world gait for both PD and HC. Based on the AUC of each gait characteristic compared to the lab, real-world gait better discriminated PD from HC, with step velocity in longer WBs (>120 s) providing the highest AUC of 0.765. In terms of PD classification, ML performance using real-world data gave better results compared to lab-based gait assessment for selected WB durations (WBs > 60 s; $30 <$ WB ≤ 60 s; > 120 s). Our findings show that testing environment and data aggregation (by WB duration) influence accuracy of ML performance and, therefore, classification of PD.

Impact of Environment and Data Aggregation on Gait Characteristics

Lab vs. Real-World

In the present study, gait assessed in the lab appeared to give different values and results compared to the real-world

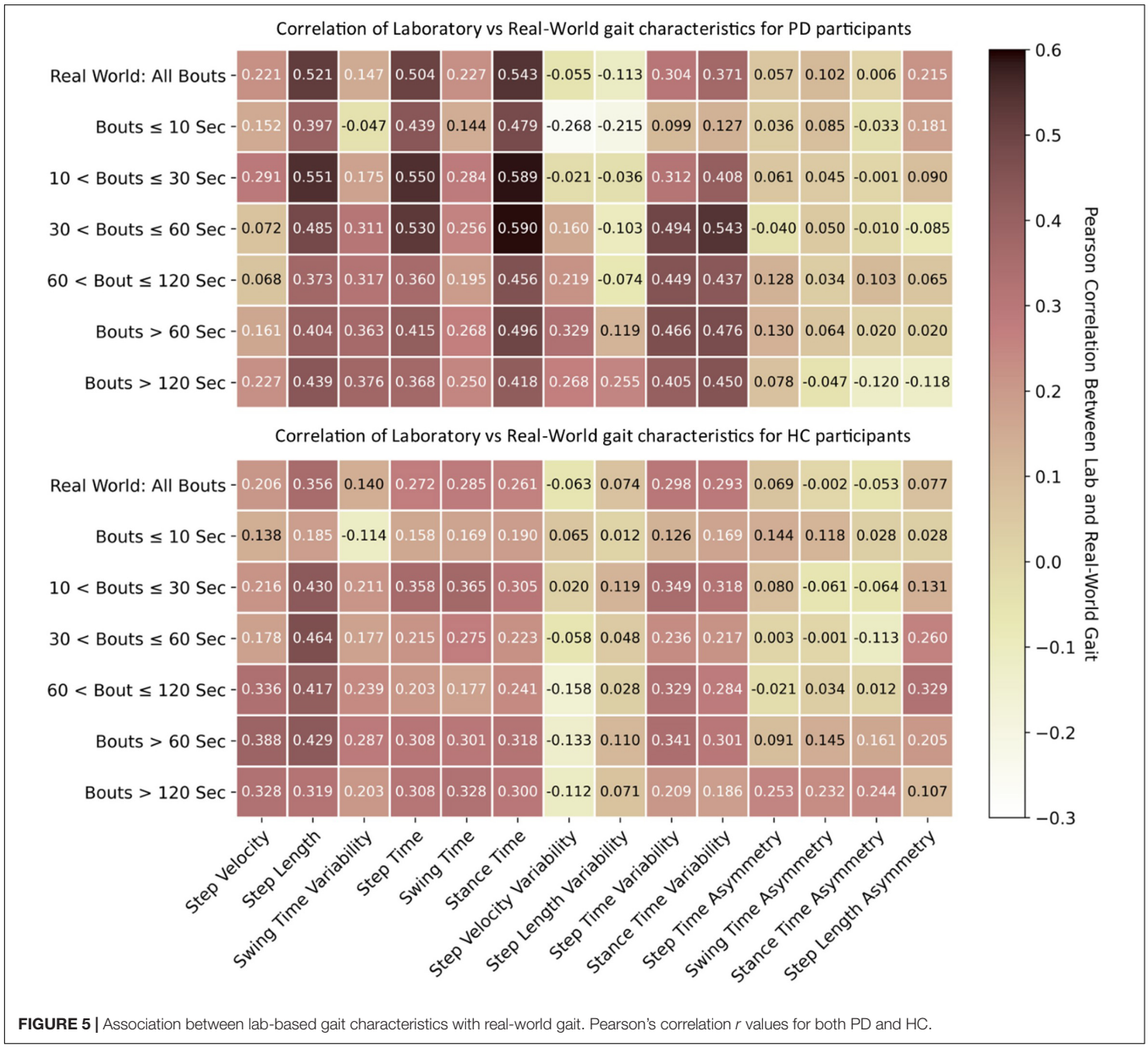


FIGURE 5 | Association between lab-based gait characteristics with real-world gait. Pearson's correlation *r* values for both PD and HC.

gait assessed for 7 consecutive days across all gait domains (pace, rhythm, variability, asymmetry, and postural control). These findings align with previous work (Del Din et al., 2016a), where PD and HC gait was assessed in the lab (10 m walk in a straight line) and in the real world for over 7 days. A major factor explaining the differences observed between environments is that in the lab, gait is measured under controlled settings during scripted tests (reflecting capacity), whilst real-world gait is characterized by natural walking behavior executed under variable settings and conditions (reflecting performance) (Del Din et al., 2016a,b, 2019; Shah et al., 2020c). People tend to walk faster with longer steps, lower variability, and higher asymmetry in the lab compared to real-world (Del Din et al., 2016a), which is evident from the present findings and in agreement with others (Takayanagi et al., 2019). Findings from this study support previous work showing that real-world gait is more variable than lab-based gait (Del Din et al., 2016a,b, 2019).

We found that the association (correlation) between lab-based and real-world gait characteristics was weak to moderate, irrespective of WB duration, suggesting real-world and lab-based gait are measuring different aspects and constructs (i.e., performance vs. capacity) of walking (Maetzler et al., 2020). These results concur with previous work showing that walking speed during a 4 m walk had low correlation with real-world gait (Van Ancum et al., 2019). One reason for the low correlation could be the heterogeneous distribution of real-world characteristics. Moreover, other gait activities, such as turning, were not accounted for. For example, people with PD tend to turn, on average, >60 times every hour (El-Gohary et al., 2014) rather than walk in perfectly

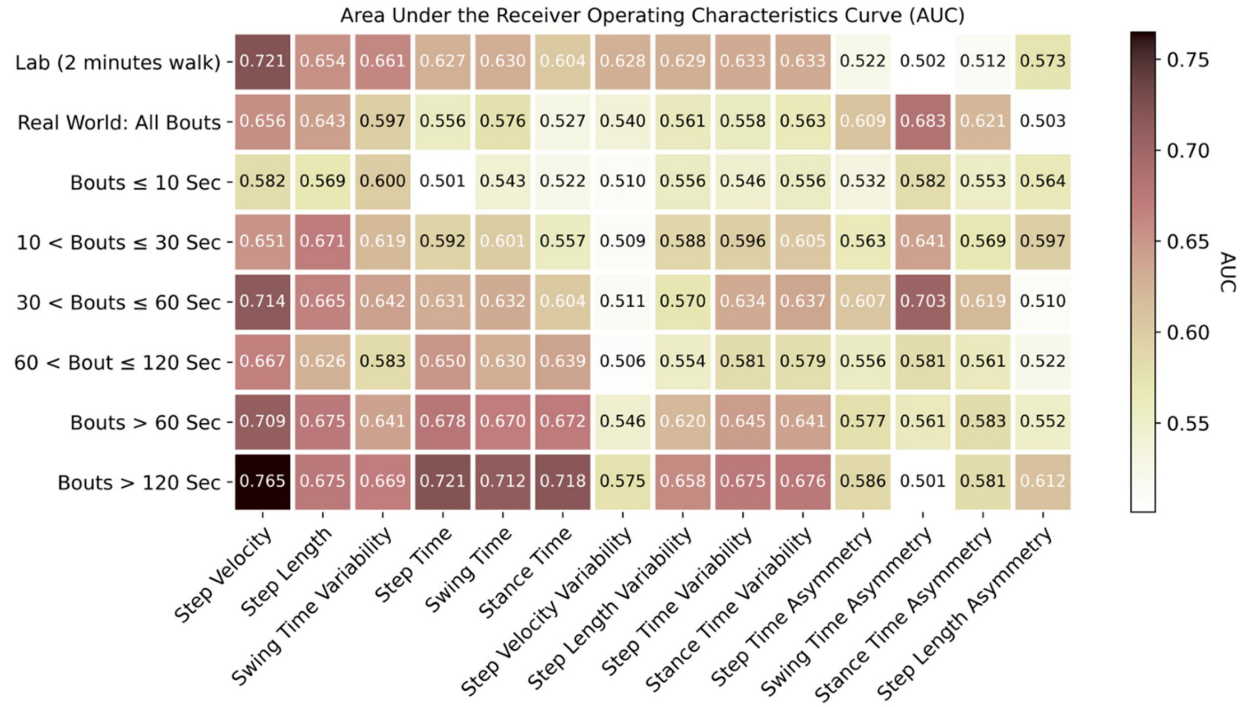


FIGURE 6 | Discrimination of PD from HC based on each individual gait characteristic with area under the receiver operating characteristics curve (AUC).

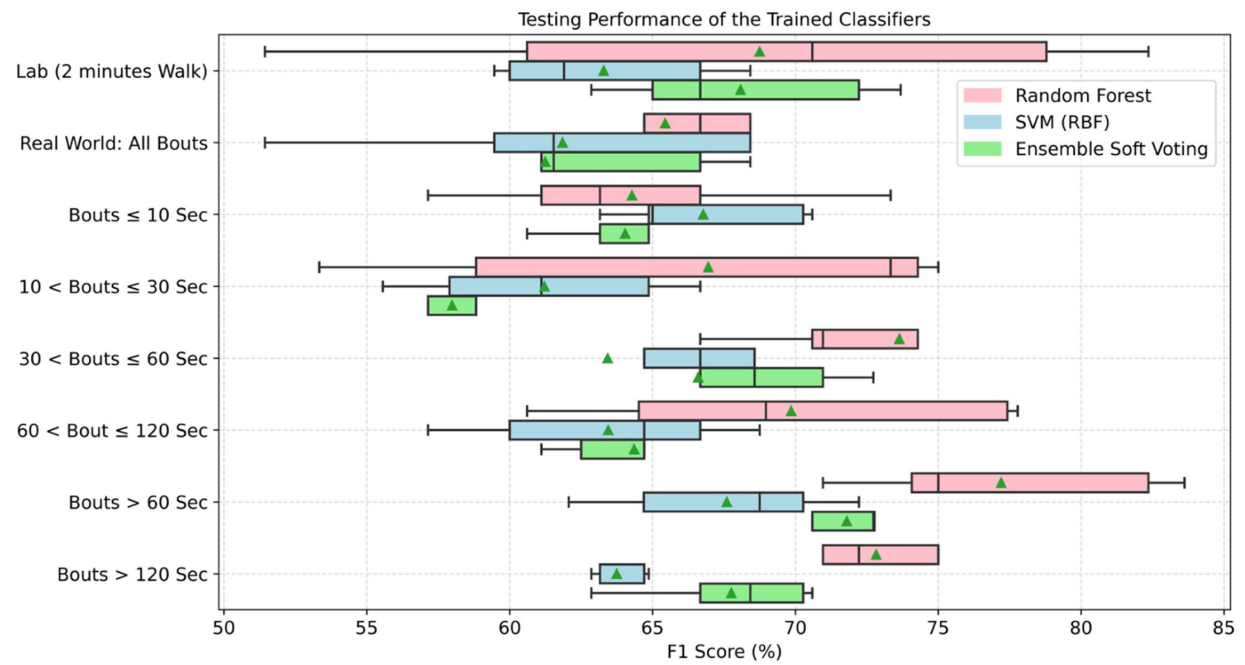


FIGURE 7 | Accuracy of trained classifiers based on the lab and real-world assessment of gait grouped by WB durations (▲ means average testing accuracy across all tests on independent dataset).

straight lines in the real-world (Galperin et al., 2019; Hillel et al., 2019), which cannot be evaluated using a tri-axial accelerometer alone.

There are many other factors influencing the complexity of real-world gait. Real-world gait is intrinsically dual task and is cognitively demanding due to complex and challenging

TABLE 3 | Evaluation metrics (accuracy, sensitivity, and specificity) of the trained classifiers on the lab and real-world test data under various walking bout (WB) durations.

Lab/WB durations	Random forest			Support vector machine			Ensemble classifier		
	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
Lab (2 min)	64.67 ± 15.02	81.62 ± 11.98	50.27 ± 19.03	51.33 ± 5.58	88.87 ± 5.56	18.12 ± 11.65	60.67 ± 5.96	88.79 ± 3.62	35.77 ± 10.71
Real-world All WB	60.67 ± 10.65	79.55 ± 15.94	43.89 ± 12.14	52.67 ± 7.23	82.95 ± 17.90	26.71 ± 5.33	53.33 ± 7.07	80.02 ± 18.85	30.72 ± 7.91
WB < 10 s	58.67 ± 9.60	78.09 ± 9.99	41.99 ± 17.87	58.67 ± 6.06	87.94 ± 9.20	32.91 ± 11.84	56.00 ± 2.79	83.75 ± 12.04	31.72 ± 7.58
10 < WB ≤ 30 s	64.67 ± 10.44	76.59 ± 15.01	54.30 ± 10.25	52.67 ± 5.96	79.41 ± 9.52	29.04 ± 6.68	52.00 ± 6.06	71.30 ± 14.78	35.43 ± 1.48
30 < WB ≤ 60 s	**72.67 ± 7.96**	**80.55 ± 11.19**	**65.35 ± 12.46**	56.67 ± 11.30	79.98 ± 15.36	36.23 ± 12.96	62.00 ± 10.95	79.69 ± 11.79	46.37 ± 16.47
60 < WB ≤ 120 s	68.00 ± 8.03	78.97 ± 10.76	58.06 ± 8.63	60.00 ± 6.24	73.87 ± 10.63	47.88 ± 8.92	60.67 ± 5.96	75.22 ± 8.75	47.81 ± 9.31
WB > 60 s	**76.27 ± 4.68**	**84.06 ± 6.36**	**70.78 ± 10.45**	**64.13 ± 2.56**	**79.85 ± 12.07**	**50.37 ± 9.98**	**68.93 ± 2.09**	**83.51 ± 6.46**	**56.03 ± 6.24**
WB > 120 s	**70.67 ± 2.79**	**83.62 ± 9.25**	**59.90 ± 10.51**	56.00 ± 2.79	82.19 ± 7.92	32.99 ± 8.49	62.00 ± 3.80	84.87 ± 8.15	41.78 ± 8.04

Bold values mean higher accuracy for PD classification.

environments in comparison to scripted gait lab tests when attention is heightened (Robles-García et al., 2015; Del Din et al., 2016a). Another important factor, especially in PD, is that medication affects gait (Ghoraani et al., 2019; Evers et al., 2020) and this is difficult to account for in the real-world where medication regimes and intake may be unknown and will impact on gait and motor fluctuations. In the present study, the lab-based gait assessment was performed one hour after medication intake in the practically defined "on" state. Therefore, we may expect to see an individual's optimal capacity. Conversely, in the real-world, walking may take place at all points of the medication cycle, resulting in on-off fluctuations in motor function and, consequently, in gait (Ghoraani et al., 2019). This can act as a confounding factor when averaging gait characteristics across different WB durations and identifies an important area for future work to understand the effect of medication on real-world gait.

Parkinson's Disease vs. Healthy Control

Gait characteristics extracted from the 2-min walk in the lab were statistically different for PD and HC compared to the real-world when combining all WBs together, except for asymmetry characteristics. These findings are difficult to compare with previous work where different protocols for gait assessment in the lab have been utilized [e.g., 10 m (Del Din et al., 2016a) or 7 m walk (Shah et al., 2020c)]. In real-world conditions, pace characteristics, such as step velocity and step length, were significantly different between PD and HC across all WBs > 10 s. Other gait characteristics (variability and asymmetry) behaved differently depending on WB duration, with differences in gait between PD and HC present for medium-to-long WB, but not for shorter WB. One possible reason for these discrepancies could be related to the algorithm, i.e., the performance of gait and step detection algorithms in shorter WBs may be challenged by noisy signals and presence of shuffling and weight transfer activities (Del Din et al., 2016a,c; Atrsaei et al., 2021). The other possible reason could be methodological: the choice of WB duration across which results are averaged may impact on gait differences, i.e., between-group gait differences found in medium WBs (which represent a high percentage of the total number of WBs) may drive results even when data are combined with results from longer WB durations as these represent a lower percentage and may offer reduced statistical power when making group comparisons (**Supplementary Figures 1, 2, 6**). Moreover, asymmetry that is quantified during shorter walking bouts in the real world may be linked to necessary gait adaptations to navigate complex environments. Results from medium duration WBs (e.g., 30–60 s) were comparable to those from longer WBs (>60 or >120 s) even though the latter represented only a small percentage [1.83% (PD) and 1.89% (HC)] of the total WBs. Gait characteristics, from every domain, were significantly different between PD and HC for medium-to-long WBs (30–60 s, >60, >120 s). These results are in line with previous work where the largest differences between the PD and HC groups were found in the longer WBs (>120 s) (Del Din et al., 2016a).

Impact of Environment and Walking Bout Duration on Parkinson's Disease Classification

Parkinson's disease (PD) classification was more accurate for lab-based gait assessment than the real-world when all the WBs were combined. Our findings for step velocity and step length yielded greater AUC, while previous work, which focused on other biomechanical characteristics, showed that foot strike angle resulted as the gait characteristic providing highest AUC (Shah et al., 2020a,c). This could be due to the different protocol (2-min walk vs. 7 min walk), cohort characteristics, and sample size (our study group mean MDS-UPDRS III score: 31; PD $n = 47$ while Shah's group mean MDS-UPDRS III score: 35 with PD $n = 29$). Discrepancies between studies show how, depending on the PD cohort disease severity, stage, sample size, and different gait characteristics may lead to higher classification performances as reflecting various level of impairment and progression. In addition, novel insights from our work showed that gait was more asymmetrical in the real-world, and this domain resulted in higher AUC than lab-based asymmetry results.

Because real-world gait presents a heterogeneous distribution, combining all WBs may increase spread of the data thereby "masking" significant differences between groups (Del Din et al., 2016a,b, 2019; Shah et al., 2020a,c,b; Atrsaei et al., 2021). As indicated previously, gait assessment conditions, such as lab and real-world, directly influence gait characteristics (Del Din et al., 2016a, 2019; Van Ancum et al., 2019; Shah et al., 2020c) with optimal walking capacity found under brief testing conditions (lab) compared to real-world performance (what people actually do during everyday activities) (Maetzler et al., 2020). ML models are directly influenced by gait features obtained under these different environments, which in turn impact classification accuracy. Therefore, combining all walking bouts obtained in the real-world can result in less optimal performance in the ML classifiers explaining our findings.

Data Aggregation by Walking Bout Duration

No previous study has investigated the impact of WB durations on the classification of PD using ML approaches. However, within univariate gait analysis, based on Del Din et al. (2016a), longer WBs were found to be better at discriminating PD from HC. This is in contrast to Shah et al. (2020a) where 90% of participants presented with WBs less than 53 strides. Therefore, only gait characteristics from short WBs $<$ 12 strides ($<$24 steps) were found to be reliable and more sensitive when discriminating PD from HC. However, due to the small sample size in Shah et al. (2020a), the effect of longer WB was possibly dampened (e.g., WBs $>$ 60 s long can have more than 53 strides) and had not been comprehensively investigated. Other factors, such as the algorithms, sensor location, and the protocol used and experimental set-up all influence the findings. In the present study, the sensor was attached on the lower back, and for Shah et al. (2020), sensors were attached to the ankles. The comprehensive approach taken in this study (i.e., quantifying various WB durations with reasonable sample size), highlighted that longer WBs were better for discriminating PD from HC.

Overall, the random forest classifier gave better classification performance as compared to SVM (Rehman et al., 2019a). ML models gave optimal performance from WBs $>$ 60 s followed by 30 $<$ WBs $\leq$ 60 s and WBs $>$ 120 s compared to lab and short WBs ($<$30 s). As discussed, real-world walking leads to various WB durations with a variety of gait speeds (Del Din et al., 2016a, 2019). In real-world conditions, both PD and HC groups performed a large number of very short WBs (e.g., $<$10 s) rather than prolonged WBs (e.g., $>$120 s). Short WBs most likely reflect habitual behaviors and moving in a constrained environment, such as a house, while longer WBs may represent walking outdoors which influence gait characteristics and ultimately the accuracy of the classifier. This is evident from the present results as shorter WBs ($<$10 s) demonstrated poor discriminative performance compared to longer ($>$60 s) WBs (accuracy of 56–59% vs. accuracy of 68–76%).

The most influential characteristics for the classification of PD were related to pace and rhythm. Particularly, step velocity, step length, swing time, stance time, swing time variability, and step time, which were identified by both random forest and SVM. The results are in line with previous work (Rehman et al., 2019a,b, 2020a) which showed that pace (step velocity and step length) are the most common and influential characteristics for not only differentiating early-stage PD in univariate fashion (e.g., t-test or with AUC), but also in ML classifiers. In addition, based on the AUC values, pace characteristics (e.g., step velocity, step length) gave optimal performance in both lab and real-world data. However, the best results were obtained in the real-world for longer WBs (30–60 s, $>$60 s, and $>$120 s). The results from this study are in line with the previous work (Shah et al., 2020a,c) where the effect of WB duration influenced the AUC, and real-world gait was found to be more sensitive for discrimination purposes.

Key Insights With Clinical Implications

- During the 2-min continuous walk in the lab, both PD and HC groups walked faster, with quicker and longer steps, lower variability, and higher asymmetry than when in the real-world, regardless of WB duration. Lab based assessment represents gait capacity, whereas real-world data reflect gait performance.
- Group differences (PD vs. HC) in gait, both in the lab and real-world conditions when combining all WBs, showed that PD participants walked slower, with shorter steps than HC. However, in the real world, significant between-group differences were influenced by WB duration (i.e., identified for WB longer than 60 or 120 s for all gait characteristics apart from those related to asymmetry). From a clinical perspective, the assessment in the clinic and outside the clinic can contain similar information. However, walking performance assessed over longer walks can offer increased sensitivity.
- Lab and real-world gait assessments assess different aspects of gait. No correlation or weak-to-moderate association was observed between the assessments. In routine clinical practice, these two streams of information can reflect different gait constructs and

therefore provide complementary information to support clinical decision making.

- Individual gait characteristics measured in the real-world and averaged across all WBs (univariate analysis) had relatively low AUC values compared to gait assessed in the lab. However, specific real-world WB durations (i.e., longer 30–60 s, >60 s, >120 s) give higher AUC compared to lab-based gait assessment. This reinforces the need to consider the impact of real-world data aggregation levels for targeting specific clinical questions/aspects (e.g., classification of PD).
- With ML-based multivariate analysis, choice of environment (lab vs. real-world) and data aggregation by WB durations clearly impacted on the ML classifier performance. Our findings suggest that ML-based models should be tested on the real-world longer WBs in clinical practice as an informed pre-screening decision-making tool for PD.
- Gait assessed with wearables in the real-world paired with ML gave reasonably accurate classification performance at early stages compared to current gold standard PD clinical diagnostics. This inexpensive and objective solution motivates its adoption in clinical practice and could be a promising addition to the current clinical diagnostic toolkit and complement clinical decision-making. An accurate early diagnosis of PD is important to ensure that timely and targeted treatments (both pharmacological and non-pharmacological) can be provided.

LIMITATIONS AND RECOMMENDATIONS FOR FUTURE WORK

There are limitations to this work. From the lab-based gait assessment, only 2 min of continuous walking were utilized and compared to $60 < \text{WBs} \leq 120$ s or $90 < \text{WBs} \leq 120$ s (**Supplementary Figures 1, 2**) of real-world gait data. However, real-world walking comprises additional complexity, with varying visual stimuli (i.e., day, night), cognitive load (single and dual task), and motor demand (i.e., uphill, downhill), which is not reflected in lab-based gait assessments. The context (e.g., indoor vs. outdoor walking) in which short and long walking bouts happen is not measured in this work. Future work is required to develop methods for characterizing contextual information. Understanding under what scenarios gait assessment could improve the classification results. In the lab, PD participants were assessed one hour after medication intake. However, in the real-world gait assessment, we could not objectively control and assess the effect of medication on gait performance. Future studies should investigate this and the effect that medication "ON" and "OFF" states could have on the results. The ML-based findings from this early PD cohort with an average disease duration of 26 months may not be generalizable to advanced PD stages. In this work, only 14 clinically relevant gait characteristics based on the heel strike and toe-off gait events were considered. In future studies, other signal-based characteristics independent of foot contact detection should be compared. Furthermore, in routine clinical practice, misdiagnosis of PD can delay subsequent intervention and treatments. Therefore, future work should look at ML classification of PD vs. atypical parkinsonian disorders (i.e., Multiple System Atrophy and Progressive Supranuclear Palsy) to identify discriminatory gait features.

CONCLUSION

In this study, we investigated the impact of environment and data aggregation by WB duration on gait characteristics and on the performance of ML models for the classification of PD. Real-world gait characteristics aggregated over medium to long WBs (e.g., WBs > 30 s) gave better discrimination performance ($0.51 \leq \text{AUC} \leq 0.77$) compared to lab-based gait characteristics ($0.51 \leq \text{AUC} \leq 0.72$), with real-world step velocity showing the highest AUC (0.77). Gait data aggregation by WB durations influenced ML classification performance. ML models applied to real-world gait showed better classification performance compared to lab data. Overall, RF trained on 14 gait characteristics aggregated over WBs > 60 s gave better performance (F1 score $= 77.20 \pm 5.51\%$) as compared to lab-based data. Findings from this study suggest that choice of environment and data aggregation by WB duration are important to achieve maximum discrimination performance and have a direct impact on ML performance for PD classification.

DATA AVAILABILITY STATEMENT

All the relevant data is either reported in the form of table or displayed in figure. The plots of all the digital gait characteristics are presented in the **Supplementary Material**. Due to data privacy and sharing agreement, raw dataset is not publically available. However, it can be made available upon reasonable request from the corresponding author. Requests to access these datasets should be directed to SD, silvia.deldin@ncl.ac.uk.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Newcastle and North Tyneside Research Ethics Committee (REC No. 09/H0906/82). Study procedures were conducted according to Declaration of Helsinki. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RR conceptualized and designed the study, and performed data analysis, statistical analysis, drafting, and critical revision of the manuscript. SD helped in conceptualization of this study, interpretation of data, and critical revision of the manuscript

for important intellectual content. YG and JS provided support for statistical analysis, interpretation, and critical revision of the manuscript for important intellectual content. LA, AY, and LR were involved in interpretation of data and critical revision of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by "Keep Control" project which is European Union's Horizon 2020 Research and Innovation ITN program under the Marie Skłodowska-Curie grant agreement no. 721577. The ICICLE-Gait study was supported by Parkinson's United Kingdom (J-0802, G-1301) and by the National Institute for Health Research (NIHR) Newcastle Biomedical Research Center (BRC) based at Newcastle upon Tyne Hospital NHS Foundation Trust and Newcastle University. This work was also supported by the NIHR/Wellcome Trust Clinical Research Facility (CRF) infrastructure at Newcastle upon Tyne Hospitals NHS Foundation Trust. All opinions are those of the authors and not the funders. Subsequent to this work, RR and SD were supported by the Innovative Medicines Initiative 2 Joint Undertaking (IMI2 JU) project IDEA-FAST–Grant Agreement 853981. In addition LR, AY, LA, and SD were also supported by the Innovative Medicines Initiative 2 Joint Undertaking (IMI2 JU) project Mobilise-D that has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement no. 820820. This JU received support from the European Union's Horizon 2020 Research and Innovation Program and the European Federation of Pharmaceutical Industries and Associations (EFPIA).

ACKNOWLEDGMENTS

We would like to thank all the participants and assessors of the ICICLE study, Rachael Lawson, and Brook Galna for their support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.808518/full#supplementary-material

REFERENCES

Atrsaei, A., Corrà, M. F., Dadashi, F., Vila-Chã, N., Maia, L., Mariani, B., et al. (2021). Gait speed in clinical and daily living assessments in Parkinson's disease patients: performance versus capacity. *NPJ Parkinsons Dis.* 7:24. doi: 10.1038/s41531-021-00171-0

Barry, G., Galna, B., Lord, S., Rochester, L., and Godfrey, A. (2015). Defining ambulatory bouts in free-living activity: impact of brief stationary periods on bout metrics. *Gait Posture* 42, 594–597. doi: 10.1016/j.gaitpost.2015.07.062

Buckley, C., Alcock, L., Mcardle, R., Rehman, R. Z. U., Del Din, S., Mazzà, C., et al. (2019). The role of movement analysis in diagnosing and monitoring neurodegenerative conditions: insights from gait and postural control. *Brain Sci.* 9:34. doi: 10.3390/brainsci9020034

De Bruin, E. D., Najafi, B., Murer, K., Uebelhart, D., and Aminian, K. (2007). Quantification of everyday motor function in a geriatric population. *J. Rehabil. Res. Dev.* 44, 417–428. doi: 10.1682/jrrd.2006.01.0003

Del Din, S., Galna, B., Godfrey, A., Bekkers, E. M., Pelosin, E., Nieuwhof, F., et al. (2019). Analysis of free-living gait in older adults with and without Parkinson's disease and with and without a history of falls: identifying generic and disease-specific characteristics. *J. Gerontol. Ser. A* 74, 500–506. doi: 10.1093/gerona/glx254

Del Din, S., Godfrey, A., Galna, B., Lord, S., and Rochester, L. (2016a). Free-living gait characteristics in ageing and Parkinson's disease: impact of environment and ambulatory bout length. *J. Neuroeng. Rehabil.* 13:46. doi: 10.1186/s12984-016-0154-5

Del Din, S., Godfrey, A., Mazzà, C., Lord, S., and Rochester, L. (2016b). Free-living monitoring of Parkinson's disease: lessons from the field. *Mov. Disord.* 31, 1293–1313. doi: 10.1002/mds.26718

Del Din, S., Godfrey, A., and Rochester, L. (2016c). Validation of an accelerometer to quantify a comprehensive battery of gait characteristics in healthy older adults and Parkinson's disease: toward clinical and at home use. *IEEE J. Biomed. Health Inf.* 20, 838–847. doi: 10.1109/JBHI.2015.2419317

Del Din, S., Hickey, A., Hurwitz, N., Mathers, J. C., Rochester, L., and Godfrey, A. (2016d). Measuring gait with an accelerometer-based wearable: influence of device location, testing protocol and age. *Physiol. Meas.* 37, 1785–1797. doi: 10.1088/0967-3334/37/10/1785

Del Din, S., Yarnall, A. J., Barber, T. R., Lo, C., Crabbe, M., Rolinski, M., et al. (2020). Continuous real-world gait monitoring in idiopathic REM sleep behavior disorder. *J. Parkinsons Dis.* 10, 283–299. doi: 10.3233/JPD-191773

Dorsey, E. R., Elbaz, A., Nichols, E., Abd-Allah, F., Abdelalim, A., Adsuar, J. C., et al. (2018). Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* 17, 939–953. doi: 10.1016/S1474-4422(18)30295-3

Dorsey, E. R., George, B. P., Leff, B., and Willis, A. W. (2013). The coming crisis: obtaining care for the growing burden of neurodegenerative conditions. *Neurology* 80, 1989–1996. doi: 10.1212/WNL.0b013e318293e2ce

El-Gohary, M., Pearson, S., Mcnames, J., Mancini, M., Horak, F., Mellone, S., et al. (2014). Continuous monitoring of turning in patients with movement disability. *Sensors* 14, 356–369. doi: 10.3390/s140100356

Emamzadeh, F. N., and Surguchov, A. (2018). Parkinson's disease: biomarkers, treatment, and risk factors. *Front. Neurosci.* 12:612. doi: 10.3389/fnins.2018.00612

Evers, L. J., Raykov, Y. P., Krijthe, J. H., De Lima, A. L. S., Badawy, R., Claes, K., et al. (2020). Real-life gait performance as a digital biomarker for motor fluctuations: the Parkinson@ home validation study. *J. Med. Internet Res.* 22:e19068. doi: 10.2196/19068

Feigin, V. L., Nichols, E., Alam, T., Bannick, M. S., Beghi, E., Blake, N., et al. (2019). Global, regional, and national burden of neurological disorders, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol.* 18, 459–480. doi: 10.1016/S1474-4422(18)30499-X

Galperin, I., Hillel, I., Del Din, S., Bekkers, E. M., Nieuwboer, A., Abbruzzese, G., et al. (2019). Associations between daily-living physical activity and laboratory-based assessments of motor severity in patients with falls and Parkinson's disease. *Parkinsonism Relat. Disord.* 62, 85–90. doi: 10.1016/j.parkreldis.2019.01.022

Ghoraani, B., Hssayeni, M. D., Bruack, M. M., and Jimenez-Shahed, J. (2019). Multilevel features for sensor-based assessment of motor fluctuation in Parkinson's disease subjects. *IEEE J. Biomed. Health Inf.* 24, 1284–1295. doi: 10.1109/JBHI.2019.2943866

Godfrey, A., Del Din, S., Barry, G., Mathers, J. C., and Rochester, L. (2015). Instrumenting gait with an accelerometer: a system and algorithm examination. *Med. Eng. Phys.* 37, 400–407. doi: 10.1016/j.medengphy.2015.02.003

Godfrey, A., Lord, S., Galna, B., Mathers, J. C., Burn, D. J., and Rochester, L. (2014). The association between retirement and age on physical activity in older adults. *Age Ageing* 43, 386–393. doi: 10.1093/ageing/aft168

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., et al. (2008). Movement disorder society-sponsored revision of

the unified Parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov. Disord.* 23, 2129–2170. doi: 10.1002/mds.22340

González, R. C., López, A. M., Rodriguez-Uría, J., Alvarez, D., and Alvarez, J. C. (2010). Real-time gait event detection for normal subjects from lower trunk accelerations. *Gait Posture* 31, 322–325. doi: 10.1016/j.gaitpost.2009.11.014

Hickey, A., Del Din, S., Rochester, L., and Godfrey, A. (2016). Detecting free-living steps and walking bouts: validating an algorithm for macro gait analysis. *Physiol. Meas.* 38, N1–N15. doi: 10.1088/1361-6579/38/1/N1

Hillel, I., Gazit, E., Nieuwboer, A., Avanzino, L., Rochester, L., Cereatti, A., et al. (2019). Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating the gaps between gait performance in the lab and during 24/7 monitoring. *Eur. Rev. Aging Phys. Act.* 16:6. doi: 10.1186/s11556-019-0214-5

Hobert, M. A., Nussbaum, S., Heger, T., Berg, D., Maetzler, W., and Heinzel, S. (2019). Progressive gait deficits in Parkinson's disease: a wearable-based biannual 5-year prospective study. *Front. Aging Neurosci.* 11:22. doi: 10.3389/fnagi.2019.00022

Hoehn, M. M., and Yahr, M. D. (1998). Parkinsonism: onset, progression, and mortality. *Neurology* 50:318. doi: 10.1212/wnl.50.2.318

Jankovic, J. (2008). Parkinson's disease: clinical features and diagnosis. *J. Neurol. Neurosurg. Psychiatry* 79, 368–376.

Khoo, T. K., Yarnall, A. J., Duncan, G. W., Coleman, S., O'brien, J. T., Brooks, D. J., et al. (2013). The spectrum of nonmotor symptoms in early Parkinson disease. *Neurology* 80, 276–281. doi: 10.1212/WNL.0b013e31827deb74

Lawson, R. A., Yarnall, A. J., Duncan, G. W., Breen, D. P., Khoo, T. K., Williams-Gray, C. H., et al. (2016). Cognitive decline and quality of life in incident Parkinson's disease: the role of attention. *Parkinsonism Relat. Disord.* 27, 47–53. doi: 10.1016/j.parkreldis.2016.04.009

Lord, S., Galna, B., Verghese, J., Coleman, S., Burn, D., and Rochester, L. (2013). Independent domains of gait in older adults and associated motor and nonmotor attributes: validation of a factor analysis approach. *J. Gerontol. Ser. A* 68, 820–827. doi: 10.1093/gerona/gls255

Maetzler, W., Domingos, J., Srulijes, K., Ferreira, J. J., and Bloem, B. R. (2013). Quantitative wearable sensors for objective assessment of Parkinson's disease. *Mov. Disord.* 28, 1628–1637. doi: 10.1002/mds.25628

Maetzler, W., Rochester, L., Bhidayasiri, R., Espay, A. J., Sánchez-Ferro, A., and Van Uem, J. M. (2020). Modernizing daily function assessment in Parkinson's disease using capacity, perception, and performance measures. *Mov. Disord.* 36, 76–82. doi: 10.1002/mds.28377

Mancini, M., King, L., Salarian, A., Holmstrom, L., Mcnames, J., and Horak, F. B. (2011). Mobility lab to assess balance and gait with synchronized body-worn sensors. *J. Bioeng. Biomed. Sci.* 12(Suppl. 1):7. doi: 10.4172/2155-9538.S1-007

McCamley, J., Donati, M., Grimpampi, E., and Mazza, C. (2012). An enhanced estimate of initial contact and final contact instants of time using lower trunk inertial sensor data. *Gait Posture* 36, 316–318. doi: 10.1016/j.gaitpost.2012.02.019

Moe-Nilssen, R. (1998). A new method for evaluating motor control in gait under real-life environmental conditions. Part 2: gait analysis. *Clin. Biomech.* 13, 328–335. doi: 10.1016/s0268-0033(98)00090-4

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., et al. (2005). The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J. Am. Geriatr. Soc.* 53, 695–699. doi: 10.1111/j.1532-5415.2005.53221.x

Nutt, J. G., and Wooten, G. F. (2005). Diagnosis and initial management of Parkinson's disease. *N. Engl. J. Med.* 353, 1021–1027.

Orendurff, M. S., Schoen, J. A., Bernatz, G. C., Segal, A. D., and Klute, G. K. (2008). How humans walk: bout duration, steps per bout, and rest duration. *J. Rehabil. Res. Dev.* 45, 1077–1089. doi: 10.1682/jrrd.2007.11.0197

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830. doi: 10.1080/13696998.2019.1666854

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ* 316, 1236–1238. doi: 10.1136/bmj.316.7139.1236

Postuma, R. B., Berg, D., Stern, M., Poewe, W., Olanow, C. W., Oertel, W., et al. (2015). MDS clinical diagnostic criteria for Parkinson's disease. *Mov. Disord.* 30, 1591–1601. doi: 10.1002/mds.26424

Powell, L. E., and Myers, A. M. (1995). The activities-specific balance confidence (ABC) scale. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* 50, M28–M34.

Przedborski, S., Vila, M., and Jackson-Lewis, V. (2003). Series introduction: neurodegeneration: what is it and where are we? *J. Clin. Invest.* 111, 3–10. doi: 10.1089/brain.2021.0059

Rehman, R. Z. U., Del Din, S., Guan, Y., Yarnall, A. J., Shi, J. Q., and Rochester, L. (2019a). Selecting clinically relevant gait characteristics for classification of early Parkinson's disease: a comprehensive machine learning approach. *Sci. Rep.* 9:17269. doi: 10.1038/s41598-019-53656-7

Rehman, R. Z. U., Del Din, S., Shi, J. Q., Galna, B., Lord, S., Yarnall, A. J., et al. (2019b). Comparison of walking protocols and gait assessment systems for machine learning-based classification of Parkinson's disease. *Sensors* 19:5363. doi: 10.3390/s19245363

Rehman, R. Z. U., Buckley, C., Micó-Amigo, M. E., Kirk, C., Dunne-Willows, M., Mazzà, C., et al. (2020a). Accelerometry-based digital gait characteristics for classification of Parkinson's disease: what counts? *IEEE Open J. Eng. Med. Biol.* 1, 65–73.

Rehman, R. Z. U., Klocke, P., Hryniv, S., Galna, B., Rochester, L., Del Din, S., et al. (2020b). Turning detection during gait: algorithm validation and influence of sensor location and turning characteristics in the classification of Parkinson's disease. *Sensors* 20:5377. doi: 10.3390/s20185377

Rizzo, G., Copetti, M., Arcuti, S., Martino, D., Fontana, A., and Logroscino, G. (2016). Accuracy of clinical diagnosis of Parkinson disease. *Neurology* 86, 566–576. doi: 10.1212/wnl.0000000000002350

Robles-García, V., Corral-Bergantiños, Y., Espinosa, N., Jácome, M. A., García-Sancho, C., Cudeiro, J., et al. (2015). Spatiotemporal gait patterns during overt and covert evaluation in patients with Parkinson's disease and healthy subjects: is there a Hawthorne effect? *J. Appl. Biomech.* 31, 189–194. doi: 10.1123/jab.2013-0319

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology* 1, 43–46.

Shah, V. V., Mcnames, J., Harker, G., Mancini, M., Carlson-Kuhta, P., Nutt, J. G., et al. (2020a). Effect of bout length on gait measures in people with and without Parkinson's disease during daily life. *Sensors* 20:5769. doi: 10.3390/s20205769

Shah, V. V., Mcnames, J., Mancini, M., Carlson-Kuhta, P., Nutt, J. G., El-Gohary, M., et al. (2020b). Digital biomarkers of mobility in Parkinson's disease during daily living. *J. Parkinsons Dis.* 10, 1099–1111. doi: 10.3233/JPD-201914

Shah, V. V., Mcnames, J., Mancini, M., Carlson-Kuhta, P., Spain, R. I., Nutt, J. G., et al. (2020c). Laboratory versus daily life gait characteristics in patients with multiple sclerosis, Parkinson's disease, and matched controls. *J. Neuroeng. Rehabil.* 17:159. doi: 10.1186/s12984-020-00781-4

Takayanagi, N., Sudo, M., Yamashiro, Y., Lee, S., Kobayashi, Y., Niki, Y., et al. (2019). Relationship between daily and in-laboratory gait speed among healthy community-dwelling older adults. *Sci. Rep.* 9:3496. doi: 10.1038/s41598-019-39695-0

Tomlinson, C. L., Stowe, R., Patel, S., Rick, C., Gray, R., and Clarke, C. E. (2010). Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov. Disord.* 25, 2649–2653. doi: 10.1002/mds.23429

Van Ancum, J. M., Van Schooten, K. S., Jonkman, N. H., Huijben, B., Van Lummel, R. C., Meskers, C. G., et al. (2019). Gait speed assessed by a 4-m walk test is not representative of daily-life gait speed in community-dwelling adults. *Maturitas* 121, 28–34. doi: 10.1016/j.maturitas.2018.12.008

Viceconti, M., Hernandez Penna, S., Dartee, W., Mazzà, C., Caulfield, B., Becker, C., et al. (2020). Toward a regulatory qualification of real-world mobility performance biomarkers in Parkinson's patients using digital mobility outcomes. *Sensors* 20:5920. doi: 10.3390/s20205920

von Campenhausen, S., Winter, Y., Rodrigues e Silva, A., Sampaio, C., Ruzicka, E., Barone, P., et al. (2011). Costs of illness and care in Parkinson's disease: an evaluation in six countries. *Eur. Neuropsychopharmacol.* 21, 180–191. doi: 10.1016/j.euroneuro.2010.08.002

Warmerdam, E., Hausdorff, J. M., Atrsaei, A., Zhou, Y., Mirelman, A., Aminian, K., et al. (2020). Long-term unsupervised mobility assessment in

movement disorders. *Lancet Neurol.* 19, 462–470. doi: 10.1016/S1474-4422(19)30397-7

Weiss, A., Brozgol, M., Dorfman, M., Herman, T., Shema, S., Giladi, N., et al. (2013). Does the evaluation of gait quality during daily life provide insight into fall risk? A novel approach using 3-day accelerometer recordings. *Neurorehabil. Neural Repair* 27, 742–752. doi: 10.1177/1545968313491004

Weiss, A., Herman, T., Giladi, N., and Hausdorff, J. M. (2014). Objective assessment of fall risk in Parkinson's disease using a body-fixed sensor worn for 3 days. *PLoS One* 9:e96675. doi: 10.1371/journal.pone.0096675

Yarnall, A. J., Breen, D. P., Duncan, G. W., Khoo, T. K., Coleman, S. Y., Firbank, M. J., et al. (2014). Characterizing mild cognitive impairment in incident Parkinson disease: the ICICLE-PD study. *Neurology* 82, 308–316. doi: 10.1212/WNL.0000000000000066

Zijlstra, A., and Zijlstra, W. (2013). Trunk-acceleration based assessment of gait parameters in older persons: a comparison of reliability and validity of four inverted pendulum based estimations. *Gait Posture* 38, 940–944. doi: 10.1016/j.gaitpost.2013.04.021

Zijlstra, W., and Hof, A. L. (2003). Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture* 18, 1–10. doi: 10.1016/s0966-6362(02)00190-x

Author Disclaimer: Content in this publication reflects the authors' view and neither IMI nor the European Union, EFPIA, or any Associated Partners are responsible for any use that may be made of the information contained herein.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 27 May 2022
doi: 10.3389/fnagi.2022.844414

Association Between Physical Activity Intensity and the Risk for Depression Among Adults From the National Health and Nutrition Examination Survey 2007–2018

Donghui Yang[1†], *Ming Yang*[1,2†], *Jianjun Bai*[1], *Yudiyang Ma*[1] *and Chuanhua Yu*[1*]

[1] *Department of Epidemiology and Biostatistics, School of Public Health, Wuhan University, Wuhan, China,* [2] *Department of Medical Affairs, The First Affiliated Hospital of Xiamen University, School of Medicine, Xiamen University, Xiamen, China*

OPEN ACCESS

Edited by:
Kristy A. Nielson, Marquette University, United States

Reviewed by:
Lan Guo, Sun Yat-sen University, China
Bruno Federico, University of Cassino, Italy

***Correspondence:**
Chuanhua Yu
yuchua@whu.edu.cn
orcid.org/0000-0002-5467-2481

[†]These authors have contributed equally to this work and share first authorship

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *28 December 2021*
Accepted: *21 April 2022*
Published: *27 May 2022*

Citation:
Yang D, Yang M, Bai J, Ma Y and Yu C (2022) Association Between Physical Activity Intensity and the Risk for Depression Among Adults From the National Health and Nutrition Examination Survey 2007–2018. Front. Aging Neurosci. 14:844414. doi: 10.3389/fnagi.2022.844414

Objective: Whether vigorous physical activities (VPA) bring additional benefits to depression prevention in comparison with moderate physical activity (MPA) remains unclear. The aim of this study was to find the correlation between the proportion of VPA to moderate-to-VPA (MVPA) (a combination of VPA and MPA) and the risk for depression, as well as to explore whether correlations differ among subgroups separated by age and sex.

Methods: The data originating from the National Health and Nutrition Examination Survey (NHANES) 2007–2018 were applied. The total amount of PA per week was obtained by multiplying frequency and duration. The proportion of VPA to MVPA was obtained among the participants who performed any MVPA. Depression was set for those who scored 10 and above in the Patient Health Questionnaire-9 (PHQ-9). The odds ratios (ORs) and 95% confidence intervals (95% CIs) for depression were evaluated using logistic regression.

Results: Among 26,849 participants of this study, only 12,939 adults were found with any MVPA, in which 748 participants with depression were detected. Logistic regression was conducted among 12,939 participants. The participants with higher than 66.7–100% of MVPA as VPA were inversely correlated with a 30% (OR = 0.70, 95% CI = 0.50, 0.99) lower risk for depression. The subgroup analyses revealed that significant correlations were only found in men and those aged 45 years and above.

Conclusion: This study suggested that a higher proportion of VPA to MVPA might be correlated with a lower risk for depression in men and those aged 45 years and above. Besides the recommendation, adults should perform 150 min MVPA per week, more time should be spent in performing VPA in MVPA among men and older adults.

Keywords: depression, vigorous physical activity (VPA), moderate-to-vigorous physical activity (MVPA), intensity, National Health and Nutrition Examination Survey (NHANES)

INTRODUCTION

Depression refers to a common and growing global mental health issue (Park and Zarate, 2019). The World Health Organization (WHO) has suggested that people with depression reach over 300 million globally, which accounts for 4.4% of the world's population (Estimates, 2017). The estimated lifetime prevalence of depression is 21% for women and 11–13% for men (Kessler et al., 2003; Belmaker and Agam, 2008). Depression generated heavy health burden studies, ranking the largest contributor to disability worldwide (Estimates, 2017). It could cause huge loss of health correlated with cardiovascular disease (Elderon and Whooley, 2013), diabetes (Roy and Lloyd, 2012), and cancer (Bortolato et al., 2017). Furthermore, depression has been reported as the major cause of suicide, and it is currently one of the top 10 causes of death in the United States (National Center for Health Statistics, United States, 2017). Thus, the onset of depression should be reduced.

Numerous studies reported risk factors for the development and progression of depression (e.g., lifestyle factors and psychophysiological and psychosocial determinants; Munoz et al., 2010; Lopresti et al., 2013; Tuithof et al., 2018). Physical activity (PA) has been reported as a vital modifiable factor for depression prevention. Several previous studies have demonstrated that PA was correlated with a lower risk for depression (Schuch et al., 2018; Dishman et al., 2021). However, benefits of depression prevention may vary with the intensity of PA. WHO recommends that adults should perform 150–300 min moderate PA (MPA) per week, 75–150 min vigorous PA (VPA) per week, or 150–300 min moderate-to-VPA (MVPA) per week, i.e., an equivalent combination of VPA and MPA (Bull et al., 2020). The assumption relating to the PA guidelines is that VPA may be associated with higher health benefits than MPA. A previous study has explored the correlation between physical intensity and depression, and it was reported that depressive symptoms increased with the intensity of PA decreasing (Lampinen et al., 2000). In addition, another study also demonstrated that VPA had the lower odds ratio (OR) for more severe depression compared with MPA (Mumba et al., 2021). Currier et al. supported the above findings and suggested that a lower risk for depression was also observed when MPA was substituted with VPA at any level (Currier et al., 2020). However, as demonstrated by a study based on the Australian Longitudinal Study on Women's Health, performing VPA did not bring a significant additional benefit to depression, except at a very high level of PA (Pavey et al., 2013). Although a considerable number of studies have investigated the association between PA and the risk for depression, it remained unclear whether VPA offered additional benefits than MPA.

Accordingly, to verify whether VPA provides additional benefits, the proportion of VPA to MVPA was calculated, and the correlation between the proportion and the risk for depression was investigated based on data from the National Health and Nutrition Examination Survey (NHANES) 2007–2018. Furthermore, it was explored whether correlations differ between age and sex in this study.

MATERIALS AND METHODS

Study Population

Data of this study originated from NHANES conducted by the National Center for Health Statistics (NCHS) of the CDC. In brief, NHANES was a nationally representative cross-sectional survey and used a multistage probability sampling design to collect information regarding health and nutritional status in the US. More detailed contents of NHANES have been previously published elsewhere (CDC, 2017). NHANES was approved by the NCHS Research Ethics Review Board. The respective survey participant provided informed consent. Information in this study was obtained from this publicly available and deidentified NHANES database so that this study was exempt from the Institutional Review Board review.

A total of 34,525 participants had complete data regarding VPA and MPA in six cycles from 2007 to 2018 in NHANES. Among 34,525 participants, we excluded 3,510 participants without a complete estimation of depression status. Furthermore, we excluded 4,166 participants with missing data regarding marital status, educational level, family income, body mass index (BMI), smoking status, and drinking status. Finally, this study included 26,849 participants, consisting of 13,910 participants not taking MVPA and 12,939 participants taking MVPA.

Ascertainment of Depression Status

Depressive symptoms in the past 2 weeks were measured using the Patient Health Questionnaire-9 (PHQ-9), a well-validated instrument to evaluate the depression status (sensitivity: 88%; specificity: 88%; Kroenke et al., 2001). The PHQ-9 questionnaire contained nine items (i.e., anhedonia, depressed mood, sleep disturbance, fatigue, appetite changes, low self-esteem, concentration problems, psychomotor disturbances, and suicidal ideation), with the respective item scoring 0 ("not at all"), 1 ("several days"), 2 ("more than half the Days"), and 3 ("nearly every day"). Participants scoring 10 and above were considered to suffer from depression (Kroenke et al., 2001).

Evaluation of Physical Activity

Information regarding PA from 2007 to 2018 was acquired in accordance with the World Health Organization Global Physical Activity Questionnaire (Armstrong and Bull, 2006). Participants were asked the following questions: "Do you do any vigorous-intensity sports, fitness, or recreational activities that cause large increases in breathing or heart rate like running or basketball for at least 10 min continuously?" and "Do you do any moderate-intensity sports, fitness, or recreational activities that cause a small increase in breathing or heart rate such as brisk walking, bicycling, swimming, or golf for at least 10 min continuously?" When they answered yes to the respective question, further questions about the frequency and duration of PA were inquired. The frequency of MPA/VPA was measured in accordance with the question "In a typical week, on how many days do you do moderate-intensity/vigorous-intensity sports, fitness or recreational activities?" Duration of MPA/VPA was measured based on the question "How much time do you spend

doing moderate-intensity/vigorous-intensity sports, fitness or recreational activities on a typical day?"

The total amount of MPA and VPA was obtained by multiplying frequency and duration, accounting for intensity, MVPA (min/week) = MPA (min/week) + [2 × VPA (min/week)]. Among the participants performing any MVPA, we obtained the proportion of VPA to MVPA as follows: VPA × 2/MVPA × 100%. The proportion of VPA to MVPA was categorized as 0–33.3%, >33.3–66.7%, and >66.7–100.0%.

Covariates

The information regarding covariates was collected through examination and questionnaire review. The covariates consisted of demographical factors (e.g., age, sex; race/ethnicity, marital status, education level, and family income) and lifestyle factors (e.g., BMI, drinking status, and smoking status). Race/ethnicity was categorized as four groups, including Hispanic, non-Hispanic white, non-Hispanic black, and other non-Hispanic. Marital status was categorized as married or living with partner; widowed, divorced, or separated; and never married. Educational level was divided into three levels, including <high school, high school, and >high school. Family poverty to income (PIR) threshold fell into three levels, including 0.0–1.0, 1.1–3.0, and >3.0. BMI, calculated as weight (kg) divided by height (m) squared, threshold fell into three levels, including <25, 25.0–29.9, and ≥30 kg/m^2. Smoking status was divided into three groups, including never smoker, former smoker, and current smoker. Drinking status was categorized into three groups, including never drinker, former drinker, and current drinker. According to the WHO guidelines, participants were categorized into two groups whether meeting the WHO guidelines (≥150 min MVPA/week) or not.

Statistical Analysis

Impacted by the complex design in NHANES, all analyses in this study considered sample weights, clustering, and stratification. Categorized variables were expressed as frequency with weighted percentage. Logistic regression was built to evaluate the ORs and 95% confidence intervals (CIs) for the correlation between the proportion of VPA to MVPA and depression. In this study, correlations between the proportion of VPA to MVPA and the risk for depression were examined in the model adjusted for multivariate, consisting of age (<45 years and ≥45 years), sex (men and women), race/ethnicity (Hispanic, non-Hispanic white, non-Hispanic black, and other non-Hispanic), marital status (married or living with partner; widowed, divorced, or separated; and never married), educational level (<high school, high school, and >high school), PIR (0.0–1.0, 1.1–3.0, and >3.0), BMI (<25.0, 25.0–29.9, and ≥30.0 kg/m^2), smoking status (never smoker, former smoker, and current smoker), drinking status (never drinker, former drinker, and current drinker), and meeting PA guideline (yes or no). Subgroups analyses were conducted to evaluate the correlation between proportion and depression among groups separated by age (<45 years and ≥45 years) and sex (men and women).

In addition, the robustness of the results of this study was evaluated through sensitivity analyses. First, the E-value was adopted to evaluate the strength of the correlations, on the risk ratio scale, of an unmeasured confounder with both exposure and outcome needed to explain away the observed correlations (VanderWeele and Ding, 2017). Second, among the 31,015 participants (14,847 with MVPA and 16,168 without MVPA) with complete information regarding PA and depression, there were missing values for drinking status ($n = 70$; 0.2%), smoking status ($n = 303$; 1.0%), BMI ($n = 326$; 1.1%), marital status ($n = 1{,}110$; 3.6%), educational level ($n = 1{,}116$; 3.6%), and PIR ($n = 2{,}840$; 9.2%). Analysis was based on multiple imputed data regarding missing values.

Data analyses were conducted using Stata version 15. A 2-sided $P < 0.05$ was considered statistically significant.

RESULTS

A total of 26,849 adults were included in this study (13,213 men and 13,636 women), including 2,501 participants with depression and 24,348 participants without depression. **Table 1** lists the basic characteristic of participants in this study. The estimation result showed that 13,910 participants did not perform MVPA in a week, accounting for 45.47% of total participants. Among the 12,939 adults reported with any MVPA, the proportion of VPA to MVPA showed the following distributions: over half of the participants (53.26%) were found with 0 to 33.3% of VPA, 15.11% were reported to be higher than 33.3–66.7% of VPA, and 31.63% were found to be more than 66.7% of VPA, respectively. Participants who were younger, men, non-Hispanic white, and married, with a higher educational level, with a high-income level, with a normal BMI (<25.0 kg/m^2), with current alcohol assumption, and with no smoking history were found to be more likely to have a higher proportion of MVPA as VPA.

After the multivariate adjustment for sociodemographic factors and lifestyle factors in logistic regression, a higher proportion of VPA to MVPA was found to be correlated with a lower risk for depression. **Table 2** lists ORs with 95%CIs of the respective covariates. Demographic risk factors for depression consisted of women, lower educational level, and income, while marriage was indicated as a protective factor. Participants reporting over 66.7–100% of MVPA as VPA were found to be inversely correlated with the risk for depression (OR = 0.70, 95% CI = 0.50, 0.99). Furthermore, the magnitude of unmeasured confounding required to clarify this inverse correlation was 2.21 between proportions with higher than 66.7–100% of MVPA as VPA and depression (**Supplementary Figure 1**).

Furthermore, **Figure 1** presents the results in subgroups analysis stratified by age and sex. Correlations differed between age and sex in this study. When the analysis was conducted among groups separated by age, no significant correlation was found between proportion and depression among those aged less than 45 years. In contrast, among those aged 45 and above, participants performing over 66.7–100% of MVPA as VPA were found to be correlated with a 56% (OR = 0.44, 95% CI = 0.24, 0.82) lower risk for depression. For analysis among groups separated by sex, men with higher than 66.7% of MVPA as VPA were found to have a significantly lower

TABLE 1 | Basic characteristics of 26,849 participants in this study from NHANES 2007–2018.

Variables	No MVPA (%)	Proportion of VPA to MVPA		
		≥0.0 and ≤33.3 (%)	>33.3 and ≤66.7 (%)	>66.7 and ≤100.0 (%)
Total	13,910 (45.47)	7,181 (53.26)	1,846 (15.11)	2,912 (31.63)
Age, years				
<45	4,774 (38.12)	2,626 (38.65)	1,156 (61.28)	2,670 (68.10)
≥45	9,136 (61.88)	4,555 (61.35)	690 (38.72)	1,242 (31.90)
Sex				
Men	6,492 (46.11)	3,299 (44.76)	1,051 (57.75)	2,371 (57.95)
Women	7,418 (53.89)	3,882 (55.24)	795 (42.25)	1,541 (42.05)
Race/ethnicity				
Hispanic	3,791 (15.92)	1,470 (10.01)	418 (13.39)	859 (12.94)
Non-Hispanic white	5,741 (64.80)	3,503 (74.44)	793 (69.08)	1,547 (67.14)
Non-Hispanic black	3,059 (12.16)	1,331 (8.40)	363 (9.34)	892 (11.27)
Other Non-Hispanic	1,319 (7.12)	877 (7.15)	272 (8.19)	614 (8.65)
Marital status				
Married or living with partner	8,235 (62.85)	4,447 (66.57)	1,102 (64.99)	2,170 (58.84)
Widowed, divorced, or separated	3,584 (21.88)	1,647 (19.15)	263 (11.07)	501 (11.41)
Never married	2,091 (15.27)	1,087 (14.27)	481 (23.95)	1,241 (29.74)
Educational level				
<high school	4,371 (22.54)	1,190 (10.05)	183 (5.77)	399 (6.16)
High school	3,599 (27.90)	1,615 (22.48)	303 (15.21)	656 (14.57)
>high school	5,940 (49.56)	4,376 (67.47)	1,360 (79.02)	2,857 (79.26)
Family income-poverty ratio				
0.0–1.0	3,630 (18.77)	1,235 (10.92)	287 (9.92)	630 (11.19)
1.1–3.0	6,436 (41.55)	2,908 (33.53)	611 (27.07)	1,369 (28.74)
>3.0	3,844 (39.68)	3,038 (55.56)	948 (63.02)	1,913 (60.07)
BMI, kg/m^2				
<25.0	3,435 (24.08)	2,024 (28.25)	640 (35.34)	1,497 (40.41)
25.0–29.9	4,390 (30.85)	2,375 (33.31)	632 (36.35)	1,370 (35.09)
≥30.0	6,085 (45.07)	2,782 (38.44)	574 (28.31)	1,045 (24.50)
Drinking status				
Never drinker	2,163 (12.18)	939 (10.11)	147 (6.77)	366 (6.75)
Ever drinker	3,137 (19.77)	1,208 (13.86)	198 (8.80)	403 (8.61)
Current drinker	8,610 (68.06)	5,034 (76.03)	1,501 (84.42)	3,143 (84.63)
Smoking status				
Never smoker	7,048 (50.01)	3,926 (54.33)	1,188 (64.24)	2,594 (66.84)
Ever smoker	3,482 (25.06)	1,967 (28.33)	363 (21.74)	760 (20.85)
Current smoker	3,380 (24.93)	1,288 (17.33)	295 (14.02)	558 (12.31)
Meeting physical activity guideline				
No	13,910 (100.00)	3,493 (49.69)	98 (5.14)	382 (8.72)
Yes	0 (0.00)	3,688 (50.31)	1,748 (94.86)	3,530 (91.28)
Depression				
No	12,157 (88.15)	6,690 (93.62)	1,762 (96.26)	3,739 (96.43)
Yes	1,753 (11.85)	491 (6.38)	84 (3.74)	173 (3.57)

NHANES, National Health and Nutrition Examination Survey; VPA, vigorous physical activity; MVPA, moderate-to-vigorous physical activity; BMI, body mass index.
The NHANES used complex design. Weight was taken into consideration. Categorized variable was described as frequency and weighted percentage.

risk for depression, whereas no significant correlation was observed in women. OR for depression was 0.52 (95% CI = 0.31, 0.88) for men with more than 66.7–100% of MVPA as VPA.

Analysis based on multiple imputed data found a similar correlation with analysis on complete data (**Supplementary Figure 2**). Furthermore, there was statistical significance in all proportion groups.

TABLE 2 | ORs and 95% CIs for depression of the proportion of VPA to MVPA and other covariates.

Variables	PHQ-9 ≥ 10		ORs (95 CIs)
	Yes	No	
Age, years			
<45	400	6,052	1.00 (reference)
≥45	348	6,139	0.87 (0.68, 1.13)
Sex			
Men	301	6,420	1.00 (reference)
Women	447	5,771	1.54 (1.25, 1.89)
Race/ethnicity			
Hispanic	192	2,555	1.00 (reference)
Non-Hispanic white	312	5,531	0.94 (0.71, 1.26)
Non-Hispanic black	163	2,423	0.85 (0.65, 1.11)
Other Non-Hispanic	81	1,682	1.06 (0.75, 1.50)
Marital status			
Married or living with partner	323	7,396	1.00 (reference)
Widowed, divorced, or separated	224	2,187	1.82 (1.40, 2.37)
Never married	201	2,608	1.89 (1.44, 2.47)
Educational level			
<high school	194	1,578	1.00 (reference)
High school	164	2,410	0.73 (0.54, 1.00)
>high school	390	8,203	0.67 (0.48, 0.94)
Family income-poverty ratio			
0.0–1.0	251	1,901	1.00 (reference)
1.1–3.0	319	4,569	0.61 (0.49, 0.77)
>3.0	178	5,721	0.40 (0.31, 0.53)
BMI, kg/m^2			
<25.0	215	3,946	1.00 (reference)
25.0–29.9	195	4,182	0.87 (0.65, 1.18)
≥30.0	338	4,063	1.56 (1.18, 2.05)
Drinking status			
Never drinker	83	1,369	1.00 (reference)
Ever drinker	135	1,674	1.45 (0.96, 2.18)
Current drinker	530	9,148	1.10 (0.77, 1.57)
Smoking status			
Never smoker	334	7,374	1.00 (reference)
Ever smoker	181	2,909	1.36 (1.02, 1.82)
Current smoker	233	1,908	2.08 (1.58, 2.74)
Meet physical activity guideline			
No	300	3,673	1.00 (reference)
Yes	448	8,518	0.78 (0.59, 1.03)
Proportion of VPA to MVPA			
≥0.0 and ≤33.3 (%)	491	6,690	1.00 (reference)
>33.3 and ≤66.7 (%)	84	1,762	0.76 (0.55, 1.05)
>66.7 and ≤100.0 (%)	173	3,739	0.70 (0.50, 0.99)

PHQ-9, Patient Health Questionnaire-9; ORs, odds ratios; CIs, confidence internals; BMI, body mass index; VPA, vigorous physical activity; MVPA, moderate-to-vigorous physical activity.

DISCUSSION

As revealed by the analyses conducted among US adults from 2007 to 2018 by NHANES, the participants with a higher proportion of VPA to MVPA might be correlated with a lower risk for depression. The analysis stratified by age and sex suggested that a lower risk for depression was primarily found among men and those aged 45 years and above. No significant correlations were found between proportion and depression in women and those aged <45 years.

About 9.3% of the participants had depression, a proportion higher than the average estimated prevalence globally (Estimates, 2017). The risk factors for depression were also found, including women, lower income, and educational level. An existing study found risk factors for depression, which included low socioeconomic status (SES), women, and comorbid chronic medical conditions (i.e., obesity) (McCarron et al., 2021), consistent with the results of this study. A meta-analysis suggested that a low SES, ascertained by the use of proxies, such as education and income, was correlated with the risk of depression (Lorant et al., 2003). Previous studies have also confirmed that PA could generate benefits and reduce the risk of depression (Schuch et al., 2018; Dishman et al., 2021). Antidepressant mechanisms of PA proved that exercise could help reduce depression through biological mechanisms and psychosocial mechanisms (e.g., neuroplasticity, inflammation, oxidative stress, and the endocrine system in biological mechanisms and self-esteem, social support, as well as self-efficacy in psychosocial mechanisms; Kandola et al., 2019). However, in 2016, only 26% of men and 19% of women performed sufficient activities in the United States (Piercy et al., 2018). Thus, the participants without MVPA were found with a high proportion. Moreover, about 27.5% of participants did not meet the PA recommendation globally in 2016 (Guthold et al., 2018). Compared with the PA level globally, the low level of PA in this study led to a higher prevalence of depression than the average estimated prevalence worldwide.

Over the past few years, the correlation between PA level and the risk for depression has been extensively investigated, and persuasive evidence has been provided, showing that those with a higher level of PA had a lower risk for depression. More attention has been gradually paid to the intensity of PA. Several recent studies explored the correlation between PA intensity and the risk for depression. A large cross-sectional study conducted among 1.2 million US adults found that the mental health burden was reduced more significantly in those performing VPA (Chekroud et al., 2018), consistent with the results of this study. A higher proportion of VPA was inversely correlated with the lower risk for depression. There was no significance in a lower proportion, which was partly due to limited participants since a significant correlation was found in all proportion groups in multiple imputed data. Accordingly, an assumption was proposed that a higher proportion of VPA to MVPA may be confirmed in larger studies.

The protective effect of a high proportion of VPA to MVPA against depression was not confirmed in women and those aged less than 45 in this study. Currier et al. conducted the study among 13,884 Australian men and demonstrated that the respective additional hour of MPA replaced with VPA was correlated with a lower risk for depression (Currier et al., 2020). In comparison, Pavey et al. investigated the correlation

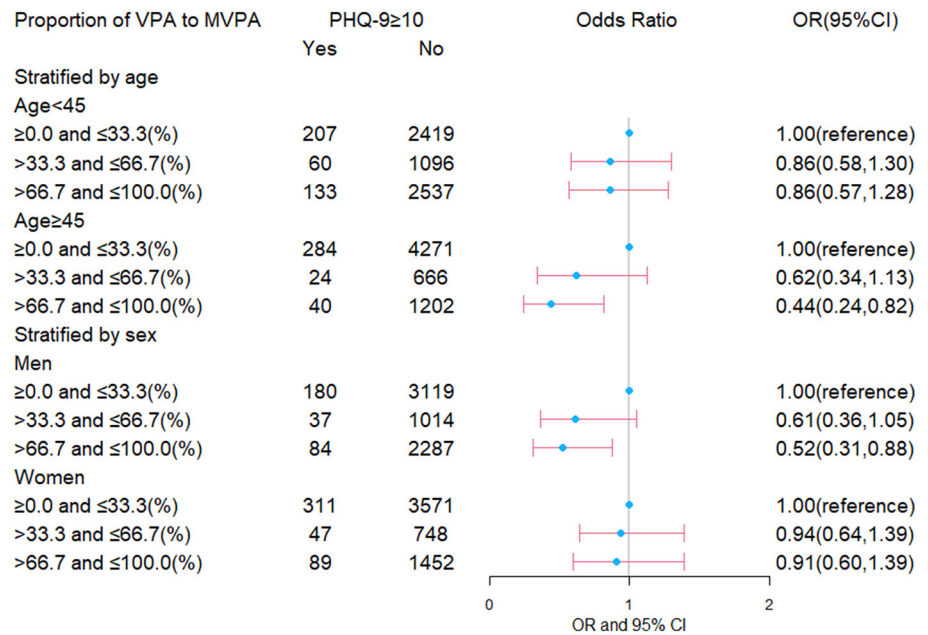


FIGURE 1 | Correlation between the proportion of VPA to MVPA and the risk for depression among groups separated by age and sex. VPA, vigorous physical activity; MVPA, moderate-to-vigorous physical activity; PHQ-9, Patient Health Questionnaire-9; OR, odds ratio; CI, confidence internal.

between PA and the risk for depression among 11,285 Australian women and found no significantly additional benefits from VPA compared with MPA except at a very high PA level (Pavey et al., 2013). The above findings were consistent with the results of this study that men with a higher proportion of VPA had a lower risk for depression, and an insignificant correlation was observed between proportion and depression in women. Women were more likely to develop episodes due to certain unique subtypes of their depression (e.g., menarche, premenstrual dysphoric disorder, postpartum depression disorder, and perimenopausal depression; Angst et al., 2002; Kessler and Bromet, 2013). Moreover, men took more VPA (Barnekow-Bergkvist et al., 1996) and women primarily did much housework (Starmer et al., 2019). VPA may be more suitable for men but not for women. A previous study has even suggested that meeting either VPA or MPA recommendation was inversely correlated with a lower risk for depression in men but not in women (Asztalos et al., 2010). It may indicate that light-intensity PA was more suitable for women to prevent from depression. Another previous study has supported our assumptions and has found significant correlations between light-intensity PA and likelihood of depression only in women and between VPA and likelihood of depression only in men (Lindwall et al., 2007).

In the analysis stratified by age, this study found a significant correlation between proportion and depression in older adults but not in younger adults. A 2-year longitudinal cohort study found that older adults were more likely to suffer from depression (Schaakxs et al., 2018). In older adults, some factors might also increase the risk for depression, including single marital status (Markkula et al., 2016), social disconnectedness (Santini et al., 2020), and less support from family and society (Wang and Zhao, 2012). Compared with younger adults, PA could prevent more chronic diseases in older adults (e.g., cancer and cardiovascular disease). Moreover, older adults' mental health also could benefit more from PA relative to younger adults. The above benefits included social interaction and engagement (Zimmer et al., 2021). Joshi et al. ever investigated the effect of quantity and type of PA on subsequent depression among old adults, and found that those performing athletic activities were at a lower risk for depression (Joshi et al., 2016). It implied that maybe VPA could benefit more in older adults. Moreover, Lampinen et al. explored the correlation between physical intensity and depression among adults aged 65 and above with 8 years of follow-up, and found that depressive symptoms increased with a decreased intensity of PA (Lampinen et al., 2000). However, the perspectives regarding participation of older people varied: for some, physical activities were not necessary and even potentially harmful; however, others were aware of the benefits of PA but reported obstacles for PA participation (Franco et al., 2015). With aging, few activities were carried out among older adults (Vancampfort et al., 2017). To decrease the OR of depression, it was necessary to change the attitude toward PA among older people and improve access to PA participation.

This study investigated the correlation between proportion and the risk for depression and verified whether differences exist between subgroups separated by age and sex. It had several strengths. This study was conducted among large general participants, which would reveal the correlation more effectively. Moreover, the correlation between proportion and depression stratified by age and sex was estimated, which could find the differences between sex and age and provided the detailed suggestions for different sex and age. In addition, the correlation

without considering the weight of intensity was analyzed besides analysis considering the weight of intensity. Furthermore, we also imputed the data and compared the analyses based on complete data and imputed data. Nonetheless, there were still some limitations in this study. Although our logistic regression model adjusted for many factors (e.g., age and sex), other confounding factors (e.g., genetic factors) were not adjusted. Furthermore, information regarding PA in this study was self-reported. Objective measurements of PA should be used to examine the volume of PA. Furthermore, this study was a cross-sectional study, which could not determine the direction of correlation or causal pathways. A randomized control trial could be implemented in further analysis.

CONCLUSION

This study implied that a higher proportion of VPA to MVPA may be correlated with a lower risk for depression. However, the above findings may be only applied to men and older adults. No significant correlation was observed in women and younger adults. Men and older adults were suggested to perform a higher proportion of VPA, while suggestion may not be suitable for women and younger adults.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.cdc.gov/nchs/nhanes/.

AUTHOR CONTRIBUTIONS

CY, DY, and MY: conception and design of the study. DY: collating data. DY and CY: analysis and/or interpretation of data. YM: visualization. DY, JB, and MY: writing the original manuscript. DY, CY, JB, YM, and MY: reviewing and editing the manuscript. CY: funding acquisition. All authors have read and agreed to the published version of the manuscript.

FUNDING

This study was funded by the National Natural Science Foundation of China (Grant Nos. 82173626 and 81773552).

ACKNOWLEDGMENTS

We thank the National Center for Health Statistics of the Centers for Disease Control and Prevention for providing the data of NHANES available on their website.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.844414/full#supplementary-material

REFERENCES

Angst, J., Gamma, A., Gastpar, M., Lepine, J. P., Mendlewicz, J., and Tylee, A. (2002). Gender differences in depression. Epidemiological findings from the European DEPRES I and II studies. *Eur. Archiv. Psychiatry Clin. Neurosci.* 252, 201–209. doi: 10.1007/s00406-002-0381-6

Armstrong, T., and Bull, F. J. (2006). Development of the world health organization global physical activity questionnaire (GPAQ). *J. Public Health* 14, 66–70. doi: 10.1007/s10389-006-0024-x

Asztalos, M., de Bourdeaudhuij, I., and Cardon, G. (2010). The relationship between physical activity and mental health varies across activity intensity levels and dimensions of mental health among women and men. *Public Health Nutr.* 13, 1207–1214. doi: 10.1017/S1368980009992825

Barnekow-Bergkvist, M., Hedberg, G., Janlert, U., and Jansson, E. (1996). Physical activity pattern in men and women at the ages of 16 and 34 and development of physical activity from adolescence to adulthood. *Scand. J. Med. Sci. Sports* 6, 359–370. doi: 10.1111/j.1600-0838.1996.tb00108.x

Belmaker, R. H., and Agam, G. (2008). Major depressive disorder. *N. Engl. J. Med.* 358, 55–68. doi: 10.1056/NEJMra073096

Bortolato, B., Hyphantis, T. N., Valpione, S., Perini, G., Maes, M., and Morris, G. (2017). Depression in cancer: the many biobehavioral pathways driving tumor progression. *Cancer Treat. Rev.* 52, 58–70. doi: 10.1016/j.ctrv.2016.11.004

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., and Cardon, G. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *Br. J. Sports Med.* 54, 1451–1462. doi: 10.1136/bjsports-2020-102955

CDC (2017). *About the National Health and Nutrition Examination Survey*. Available online at: https://www.cdc.gov/nchs/nhanes/about_nhanes.htm (accessed August 10, 2021).

Chekroud, S. R., Gueorguieva, R., Zheutlin, A. B., Paulus, M., Krumholz, H. M., and Krystal, J. H. (2018). Association between physical exercise and mental health in 1.2 million individuals in the USA between 2011 and 2015: a cross-sectional study. *Lancet Psychiatr.* 5, 739–746. doi: 10.1016/S2215-0366(18)30227-X

Currier, D., Lindner, R., Spittal, M. J., Cvetkovski, S., Pirkis, J., and English, D. R. (2020). Physical activity and depression in men: increased activity duration and intensity associated with lower likelihood of current depression. *J. Affect. Disord.* 260, 426–431. doi: 10.1016/j.jad.2019.09.061

Dishman, R. K., McDowell, C. P., and Herring, M. P. (2021). Customary physical activity and odds of depression: a systematic review and meta-analysis of 111 prospective cohort studies. *Br. J. Sports Med.* 55, 926–934. doi: 10.1136/bjsports-2020-103140

Elderon, L., and Whooley, M. A. (2013). Depression and cardiovascular disease. *Progr. Cardiovasc. Dis.* 55, 511–523. doi: 10.1016/j.pcad.2013.03.010

Estimates, G. H. (2017). *Depression and Other Common Mental Disorders*. Available online at: https://www.who.int/publications/i/item/depression-global-health-estimates (accessed January 3, 2017).

Franco, M. R., Tong, A., Howard, K., Sherrington, C., Ferreira, P. H., and Pinto, R. Z. (2015). Older people's perspectives on participation in physical activity: a systematic review and thematic synthesis of qualitative literature. *British journal of sports medicine* 49, 1268–1276. doi: 10.1136/bjsports-2014-094015

Guthold, R., Stevens, G. A., Riley, L. M., and Bull, F. C. (2018). Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1.9 million participants. *Lancet Glob. Health* 6, e1077–e86. doi: 10.1016/S2214-109X(18)30357-7

Joshi, S., Mooney, S. J., Kennedy, G. J., Benjamin, E. O., Ompad, D., and Rundle, A. G. (2016). Beyond METs: types of physical activity and depression among older adults. *Age Ageing* 45, 103–109. doi: 10.1093/ageing/afv164

Kandola, A., Ashdown-Franks, G., Hendrikse, J., Sabiston, C. M., and Stubbs, B. (2019). Physical activity and depression: towards understanding the antidepressant mechanisms of physical activity. *Neurosci. Biobehav. Rev.* 107, 525–539. doi: 10.1016/j.neubiorev.2019.09.040

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., and Merikangas, K. R. (2003). The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *J. Am. Med. Assoc.* 289, 3095–3105. doi: 10.1001/jama.289.23.3095

Kessler, R. C., and Bromet, E. J. (2013). The epidemiology of depression across cultures. *Ann. Rev. Public Health* 34, 119–138. doi: 10.1146/annurev-publhealth-031912-114409

Kroenke, K., Spitzer, R. L., and Williams, J. B. (2001). The PHQ-9: validity of a brief depression severity measure. *J. Gen. Intern. Med.* 16, 606–613. doi: 10.1046/j.1525-1497.2001.016009606.x

Lampinen, P., Heikkinen, R. L., and Ruoppila, I. (2000). Changes in intensity of physical exercise as predictors of depressive symptoms among older adults: an eight-year follow-up. *Prev. Med.* 30, 371–380. doi: 10.1006/pmed.2000.0641

Lindwall, M., Rennemark, M., Halling, A., Berglund, J., and Hassmen, P. (2007). Depression and exercise in elderly men and women: findings from the Swedish national study on aging and care. *J. Aging Phys. Act.* 15, 41–55. doi: 10.1123/japa.15.1.41

Lopresti, A. L., Hood, S. D., and Drummond, P. D. (2013). A review of lifestyle factors that contribute to important pathways associated with major depression: diet, sleep and exercise. *J. Affect. Disord.* 148, 12–27. doi: 10.1016/j.jad.2013.01.014

Lorant, V., Deliège, D., Eaton, W., Robert, A., Philippot, P., and Ansseau, M. (2003). Socioeconomic inequalities in depression: a meta-analysis. *Am. J. Epidemiol.* 157, 98–112. doi: 10.1093/aje/kwf182

Markkula, N., Harkanen, T., Nieminen, T., Pena, S., Mattila, A. K., and Koskinen, S. (2016). Prognosis of depressive disorders in the general population- results from the longitudinal Finnish Health 2011 Study. *J. Affect. Disord.* 190, 687–696. doi: 10.1016/j.jad.2015.10.043

McCarron, R. M., Shapiro, B., Rawles, J., and Luo, J. (2021). Depression. *Ann. Intern. Med.* 174, Itc65-itc80. doi: 10.7326/AITC202105180

Mumba, M. N., Nacarrow, A. F., Cody, S., Key, B. A., Wang, H., and Robb, M. (2021). Intensity and type of physical activity predicts depression in older adults. *Aging Mental Health* 25, 664–671. doi: 10.1080/13607863.2020.1711861

Munoz, R. F., Cuijpers, P., Smit, F., Barrera, A. Z., and Leykin, Y. (2010). Prevention of major depression. *Ann. Rev. Clin. Psychol.* 6, 181–212. doi: 10.1146/annurev-clinpsy-033109-132040

National Center for Health Statistics, United States (2017). *Health, United States, 2016: With Chartbook on Long-term Trends in Health*. Hyattsville, MD: National Center for Health Statistics (US).

Park, L. T., and Zarate, C. A. (2019). Depression in the primary care setting. *N. Engl. J. Med.* 380, 559–568. doi: 10.1056/NEJMcp1712493

Pavey, T. G., Peeters, G., Bauman, A. E., and Brown, W. J. (2013). Does vigorous physical activity provide additional benefits beyond those of moderate? *Med. Sci. Sports Exerc.* 45, 1948–1955. doi: 10.1249/MSS.0b013e3182940b91

Piercy, K. L., Troiano, R. P., Ballard, R. M., Carlson, S. A., Fulton, J. E., and Galuska, D. A. (2018). The physical activity guidelines for Americans. *J. Am. Med. Assoc.* 320, 2020–2028. doi: 10.1001/jama.2018.14854

Roy, T., and Lloyd, C. E. (2012). Epidemiology of depression and diabetes: a systematic review. *J. Affect. Disord.* 142(Suppl.), S8–S21. doi: 10.1016/S0165-0327(12)70004-6

Santini, Z. I., Jose, P. E., York Cornwell, E., Koyanagi, A., Nielsen, L., and Hinrichsen, C. (2020). Social disconnectedness, perceived isolation, and symptoms of depression and anxiety among older Americans (NSHAP): a longitudinal mediation analysis. *Lancet Public Health* 5, e62–e70. doi: 10.1016/S2468-2667(19)30230-0

Schaakxs, R., Comijs, H. C., Lamers, F., Kok, R. M., Beekman, A. T. F., and Penninx, B. (2018). Associations between age and the course of major depressive disorder: a 2-year longitudinal cohort study. *Lancet Psychiatry* 5, 581–590. doi: 10.1016/S2215-0366(18)30166-4

Schuch, F. B., Vancampfort, D., Firth, J., Rosenbaum, S., Ward, P. B., and Silva, E. S. (2018). Physical activity and incident depression: a meta-analysis of prospective cohort studies. *Am. J. Psychiatry* 175, 631–648. doi: 10.1176/appi.ajp.2018.17111194

Starmer, A. J., Frintner, M. P., Matos, K., Somberg, C., Freed, G., and Byrne, B. J. (2019). Gender discrepancies related to pediatrician work-life balance and household responsibilities. *Pediatrics* 144, e20182926. doi: 10.1542/peds.2018-2926

Tuithof, M., Ten Have, M., van Dorsselaer, S., Kleinjan, M., Beekman, A., de Graaf, R., et al. (2018). Course of subthreshold depression into a depressive disorder and its risk factors. *J. Affect. Disord.* 241, 206–215. doi: 10.1016/j.jad.2018.08.010

Vancampfort, D., Stubbs, B., Firth, J., Hallgren, M., Schuch, F., and Lahti, J. (2017). Physical activity correlates among 24,230 people with depression across 46 low- and middle-income countries. *J. Affect. Disord.* 221, 81–88. doi: 10.1016/j.jad.2017.06.012

VanderWeele, T. J., and Ding, P. (2017). Sensitivity analysis in observational research: introducing the *E*-value. *Ann. Intern. Med.* 167, 268–274. doi: 10.7326/M16-2607

Wang, J., and Zhao, X. (2012). Family functioning and social support for older patients with depression in an urban area of Shanghai, China. *Archiv. Gerontol. Geriatr.* 55, 574–579. doi: 10.1016/j.archger.2012.06.011

Zimmer, C., McDonough, M. H., Hewson, J., Toohey, A., Din, C., Crocker, P. R. E., et al. (2021). Experiences with social participation in group physical activity programs for older adults. *J. Sport Exerc. Psychol.* 43, 335–344. doi: 10.1123/jsep.2020-0335

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fnagi.2022.886491

The Instrumental Activities of Daily Living in Parkinson's Disease Patients Treated by Subthalamic Deep Brain Stimulation

Ondrej Bezdicek[1†‡], Josef Mana[1†‡], Filip Růžička[1], Filip Havlik[1†], Anna Fečíková[1], Tereza Uhrová[1], Evžen Růžička[1†], Dušan Urgošík[2] and Robert Jech[1†]*

[1] Department of Neurology and Centre of Clinical Neuroscience, First Faculty of Medicine and General University Hospital in Prague, Charles University, Prague, Czechia, [2] Department of Stereotactic and Radiation Neurosurgery, Na Homolce Hospital, Prague, Czechia

Background: Everyday functioning and instrumental activities of daily living (IADL) play a vital role in preserving the quality of life in patients with Parkinson's disease (PD) after deep brain stimulation of the subthalamic nucleus (STN-DBS).

Objective: The main goal of the current study was to examine IADL change in pre-and post-surgery of the STN-DBS. We also analyzed the influence of the levodopa equivalent daily dose (LEDD) and global cognitive performance (Dementia Rating Scale; DRS-2) as covariates in relation to IADL.

Methods: Thirty-two non-demented PD patients were administered before and after STN-DBS neurosurgery the Penn Parkinson's Daily Activities Questionnaire (PDAQ; self-report), the DRS-2 and Beck Depression Inventory (BDI-II) to assess IADL change, global cognition, and depression.

Results: We found a positive effect of STN-DBS on IADL in the post-surgery phase. Moreover, lower global cognition and lower LEDD are predictive of lower IADL in both pre-surgery and post-surgery examinations.

Summary/Conclusion: STN-DBS in PD is a safe method for improvement of everyday functioning and IADL. In the post-surgery phase, we show a relation of IADL to the severity of cognitive impairment in PD and to LEDD.

Keywords: activities of daily living, deep brain stimulation, cognition, everyday abilities, subthalamic nucleus

HIGHLIGHTS

- There is a significant effect of STN-DBS treatment on everyday functioning improvement 1 year after the surgery in Parkinson's disease patients.
- We found also positive effects of cognitive performance, and LEDD as well as a negative effect of depressive symptoms on everyday functioning both before and after STN-DBS surgery.
- We provide a table detailing what changes can one expect in a patient's everyday functioning depending on their pre- and post-surgery LEDD.

OPEN ACCESS

Edited by:
Fabio Pilato, Policlinico Universitario Campus Bio-Medico, Italy

Reviewed by:
Massimo Marano, Campus Bio-Medico University, Italy
Carla Piano, Agostino Gemelli University Polyclinic (IRCCS), Italy

****Correspondence:***
Ondrej Bezdicek
ondrej.bezdicek@lf1.cuni.cz

†ORCID:
Ondrej Bezdicek
orcid.org/0000-0002-5108-0181
Josef Mana
orcid.org/0000-0002-7817-3978
Filip Havlik
orcid.org/0000-0003-4442-6743
Evžen Růžička
orcid.org/0000-0002-4893-9661
Robert Jech
orcid.org/0000-0002-9732-8947

‡These authors have contributed equally to this work

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *28 February 2022*
Accepted: *30 May 2022*
Published: *17 June 2022*

Citation:
Bezdicek O, Mana J, Růžička F, Havlik F, Fečíková A, Uhrová T, Růžička E, Urgošík D and Jech R (2022) The Instrumental Activities of Daily Living in Parkinson's Disease Patients Treated by Subthalamic Deep Brain Stimulation. Front. Aging Neurosci. 14:886491. doi: 10.3389/fnagi.2022.886491

INTRODUCTION

Deficits in everyday functioning are highly associated with the evolution of Parkinson's disease (PD) (Young et al., 2010; Giovannetti et al., 2012; Altieri et al., 2021; Becker et al., 2022). Especially in clinically advanced stages when patients are suffering from deteriorating cognitive impairment such as mild cognitive impairment (PD-MCI) and later from dementia due to PD (PDD), instrumental activities of daily living (IADL) such as meal preparation, shopping, and medication management are afflicted (Pirogovsky et al., 2014; Foster and Doty, 2021). A standardized assessment of IADL can be specific for the diagnosis of PDD (Christ et al., 2013). Those PD-MCI patients who show more impaired cognitive- than motor-driven IADL have a higher hazard of conversion to PDD (Becker et al., 2020). Regarding PD motor dysfunction subtypes, *de novo* PD patients with postural instability-gait difficulties motor subtype present on average larger deterioration in IADL than those with tremor dominant subtype (Hariz and Forsgren, 2011).

Deep brain stimulation (DBS) of the subthalamic nucleus (STN-DBS) is a standard treatment for medication-refractory movement symptoms of PD (Perlmutter and Mink, 2006; Bronstein et al., 2011; Okun, 2014; Mueller et al., 2020). A number of works have shown that this therapy is highly effective in regaining control over PD motor symptoms and improving patients' quality of life (QoL), as well as in reducing the levodopa equivalent daily dose (LEDD) (Tomlinson et al., 2010; Bratsos et al., 2018; Tödt et al., 2022).

Hence DBS is a treatment option for advanced PD when the side effects of dopaminergic treatment are intolerable (Moro and Lang, 2006; Bronstein et al., 2011) and there are the differential effects of levodopa vs. DBS on brain motor activity (Mueller et al., 2020), a relation between LEDD and IADL in PD in pre- and post-surgery should come under closer scrutiny.

The effect on QoL is stable in time in a 1-year perspective and the IADL performance correlates with the improvement in QoL (Gorecka-Mazur et al., 2019). On the contrary, activity limitations are the strongest predictor of QoL (Soh et al., 2013). However, the improvement after 12 months after the surgery was noticeable only in some IADLs, especially in shopping and food preparation (Gorecka-Mazur et al., 2019).

Previous research of IADL in PD patients after STN-DBS used general IADL scales, such as Lawton's IADL scale, to investigate everyday performance (Gorecka-Mazur et al., 2019) or the Unified Parkinson's Disease Rating Scale (UPDRS part II) or the Schwab and England Activities of Daily Living (Deuschl et al., 2006; Kleiner-Fisman et al., 2010; Odekerken et al., 2013; Jiang et al., 2015; Tödt et al., 2022). In the current study, we sought to use a disease-specific questionnaire developed to assess IADLs in PD (Brennan et al., 2016a). The Penn Parkinson's Daily Activities Questionnaire (PDAQ) shows a high discriminant validity between PD with normal cognition in comparison to PD-MCI and PDD (Brennan et al., 2016a).

Cognitive impairment is a core non-motor feature of Parkinson's disease (PD) and a major source of disability in PDD (Cahn et al., 1998; Bronnick et al., 2006). Especially, impaired global cognition and deficits in attention and visual memory are the most predictive of developing a PDD (Lawson et al., 2021). PD-MCI as a pre-dementia phase and PDD both represent risk factors that are associated with poorer everyday functioning and IADL (Litvan et al., 2012; Martin et al., 2013; Pirogovsky et al., 2013; Pirogovsky et al., 2014; Hoogland et al., 2017; Schmitter-Edgecombe et al., 2021).

Thus, the principal aim of the study was to assess post-surgery change (i.e., decline or improvement) in self-reported IADL in relation to the pre-surgery evaluation. Second, we aimed to outline the relationship of IADL to dopaminergic medication, depressive symptoms and cognitive performance in PD patients treated with STN DBS.

MATERIALS AND METHODS

Participants

Parkinson's disease patients were recruited from the Movement Disorders Center, Department of Neurology, First Faculty of Medicine and General University Hospital in Prague. All patients were examined by a neurologist specializing in movement disorders and met the UK PD Society Brain Bank criteria (Hughes et al., 1992). All of them were suffering from motor fluctuations and/or disabling dyskinesias and were indicated for treatment with STN DBS (demographic and clinical details in **Table 1**). Exclusion criteria were as follows: PD dementia according to MDS criteria (Emre et al., 2007), atypical or secondary parkinsonism, severe or moderate depression according to Beck Depression Inventory (BDI-II) and psychiatric evaluation, florid psychotic manifestations (hallucinations or delusions), anticholinergic medications and other medical or neurological conditions potentially resulting in cognitive impairment (e.g., epileptic seizure, tumor, stroke, or head trauma). All PD patients were under dopaminergic therapy (i.e., levodopa, dopamine agonist, or a combination of them), and levodopa's equivalent daily dose for each patient was calculated before and after surgery (Tomlinson et al., 2010). Bilateral STN DBS implantation was performed as previously described (Jech et al., 2006; Urgosik et al., 2011; Jech et al., 2012). STN DBS parameters are reported in **Table 1**. A total of 32 PD patients (mean age 55.5 $\pm$ 7.8 years pre-surgery, 56% males) participated in the study. Patients were assessed before (4.9 $\pm$ 5.6 months) and 1 year after the surgery (12.4 $\pm$ 0.9 months). All patients gave their written informed consent for participation. The study was approved by the Ethics Committee of the General University Hospital in Prague, Czechia.

Assessments

Neuropsychological Examination

All patients underwent a comprehensive and recommended pre-surgery (pre-test) evaluation including neuropsychological, psychiatric, and neurological examinations by a trained movement disorders specialist in each field (Kubu, 2018). The patients were followed up in a post-surgery (post-test) 1 year after the neurosurgery with the identical protocol (mean retest interval 12.4 $\pm$ 0.9 months). The pre-surgery neuropsychological assessment was performed with regular dopaminergic therapy

TABLE 1 | Demographic, clinical, and cognitive characteristics of the sample (N = 32).

	Pre-surgery	Post-surgery
Age (years)	55.50 ± 7.78	56.95 ± 7.79
Education (years)	14.20 ± 3.25	–
Sex (males)	18 (56%)	–
Disease duration at surgery (years)	11.37 ± 3.67	–
LEDD (mg)	1819.77 ± 693.73	833.32 ± 498.48
Levodopa test (% response)	58.42 ± 11.79	–
MDS-UPDRS III (medication ON)	18.76 ± 9.13	–
MDS-UPDRS III (medication OFF)	44.12 ± 15.05	–
MDS-UPDRS III (stimulation ON)*	–	26.25 ± 10.00
MDS-UPDRS III (stimulation OFF)*	–	45.16 ± 14.04
PDAQ-15 (range 0–60)	51.34 ± 7.49	52.34 ± 6.35
DRS-2 (range 0–144)	139.28 ± 3.62	139.44 ± 3.33
BDI-II (range 0–63)	10.38 ± 7.20	9.91 ± 6.90
Stimulation parameters		
Current right (mA)	–	2.24 ± 0.55
Current left (mA)	–	2.21 ± 0.60
Pulse duration right (μs)	–	62.81 ± 8.88
Pulse duration left (μs)	–	63.64 ± 9.94
Frequency right (Hz)	–	129.06 ± 18.38
Frequency left (Hz)	–	125.76 ± 11.73

**Post-surgery MDS-UPDRS III testing was done in the OFF mediation condition; μs, microseconds; BDI-II, Beck Depression Rating Scale, second edition; DRS-2, Dementia Rating Scale, second edition; Hz, Hertz; LEDD, levodopa equivalent daily dose; mA, milliamperes; MDS-UPDRS III, Movement Disorder Society Unified Parkinson's Disease Rating Scale, motor part; PDAQ-15, The Penn Parkinson's Daily Activities Questionnaire-15. The values are presented in a format mean ± standard deviation or number of observations (percentage from the whole sample).*

(ON medication), in the post-surgery phase, patients were examined in both STN DBS ON with optimal stimulation parameters and the ON medication condition.

The neuropsychological assessment in pre-test–post-test followed the standard Movement Disorder Society neuropsychological battery at Level I for PD-MCI (Litvan et al., 2012; Bezdicek et al., 2016; Bezdicek et al., 2017): the cognitive performance was assessed by Mattis Dementia Rating Scale, second edition (DRS-2) (Jurica et al., 2001; Bezdicek et al., 2015). The IADLs and everyday functioning were measured by the PDAQ self-report (Shulman et al., 2016). The PDAQ brief version is an item-response theory (IRT)-based questionnaire consisting of 15 items, showing very good psychometric properties that were developed specifically for IADL deficits in PD (Brennan et al., 2016a,b). Finally, depressive symptoms were assessed with the Beck Depression Scale, second edition (BDI-II) (Beck et al., 1996; Ciharova et al., 2020).

Neurological and Psychiatric Examination

All patients underwent a comprehensive clinical evaluation that included medical history, medication status, and motor status by the Movement Disorders Society Unified Parkinson's Disease Rating Scale, part three (MDS-UPDRS-III). Scores of patients who underwent the older version of the Unified Parkinson's Disease Rating Scale (UPDRS-III) were converted to the MDS-UPDRS III scale using the method described by Hentz et al. (2015). All PD patients were treated with dopaminergic therapy, consisting of levodopa, dopamine agonists or a combination of them, and assessed in medication ON. Four days before the patient's visit, dopamine agonists were substituted with equivalent doses of levodopa. The LEDD was calculated at each assessment time-point according to Tomlinson et al. (2010).

A psychiatric evaluation was done before the surgery to exclude pre-psychotic or florid psychotic symptoms or mood disorders including suicidal thoughts or any other potential risky neuropsychiatric complications after the neurosurgery (Foley et al., 2018).

Causal Assumptions

Our causal assumptions are represented in the form of a directed acyclic graph (DAG) depicted in **Figure 1**. Representing study design *via* a DAG offers several benefits including serving as an explicit statement of causal assumptions that can be questioned by other researchers, providing a framework for the interpretation of results and indicating which covariates should be controlled during the analysis (Pearl, 2009; McElreath, 2020). In short, in the current study, we assume that clinicians based their decision on whether to treat a PD patient with STN DBS in part of the patient's preoperative IADL, objective cognitive performance and LEDD. While the level of depressive symptoms as assessed by BDI-II is not directly considered for STN DBS treatment in our center, patients with the depressive syndrome as assessed by an independent psychiatric evaluation are rejected for STN DBS due to possible suicidal attempts, and since the depressive symptom also likely leads to high BDI-II score, we assume a common cause relationship between pre-surgery BDI-II and STN DBS surgery as indicated by the dashed double-headed arrow in **Figure 1**. On the other hand, DBS treatment itself is assumed to influence postoperative IADL, objective cognitive performance, level of depressive symptoms and LEDD. Cross-sectionally (i.e., either pre-or post-surgery), LEDD is assumed to influence depressive symptoms directly and objective cognitive performance and IADL both directly as well as indirectly *via* the effect of depressive symptoms. Finally, because IADL was repeatedly assessed by an IRT-based PDAQ questionnaire we expect item- and patient-specific effects on the outcomes.

Since our sample contains only patients treated with STN DBS and no control group, we were not able to estimate either the total or direct effect of DBS on IADL. However, based on the DAG in **Figure 1**, we can estimate the direct effects of objective cognitive performance, depressive symptoms and LEDD on IADL as well as direct post-surgery change in IADL in STN DBS patients by controlling for LEDD, objective cognitive performance and depressive symptoms as well as item- and patient-specific effects.

Statistical Analysis

Following the implications of DAG presented in **Figure 1**, the data were analyzed by a generalized linear mixed model (GLMM) with responses to each item of PDAQ as an outcome, the time of assessment (pre- vs. post-surgery), LEDD, DRS-2, and BDI-II as fixed effects and item- and patient-specific random effects. Interactions of the time of assessment and all

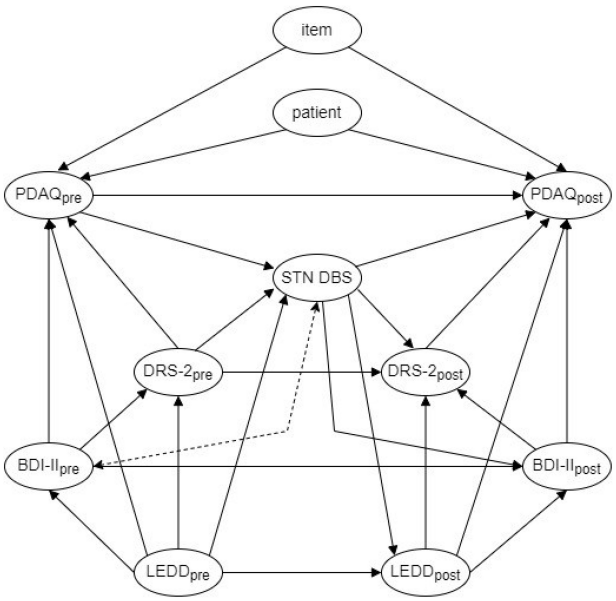


FIGURE 1 | A directed acyclic graph representing causal assumptions of the relationships between included variables. STN DBS, subthalamic nucleus deep brain stimulation; BDI-II, Beck Depression Inventory before DBS treatment (BDI-II_{pre}) and after DBS treatment (BDI-II_{post}); DRS-2, Dementia Rating Scale, second edition before DBS treatment (DRS-2_{pre}) and after DBS treatment (DRS-2_{post}); LEDD, levodopa equivalent daily dose before DBS treatment ($LEDD_{pre}$) and after DBS treatment ($LEDD_{post}$); PDAQ, The Penn Parkinson's Daily Activities Questionnaire before DBS treatment ($PDAQ_{pre}$) and after DBS treatment ($PDAQ_{post}$). STN DBS was considered to be adjusted for in each of our analyses due to the lack of a control group. Dashed double arrow between BDI-II_{pre} and STN DBS indicates a common cause assumption—this is because even though BDI-II is not used directly to decide whether patients receive STN DBS in our center, patients with clinical depression according to an independent psychiatric evaluation are both rejected to STN DBS and at risk of high BDI-II.

LEDD (medication), DRS-2 (cognitive performance), and BDI-II (depressive symptoms) were also included and modeled as fixed effects to explore whether the effect of the latter three variables on IADL changes after as compared to before STN DBS surgery. Since PDAQ consists of 15 items scored by the patient's self-reported difficulty in performing each specific IADL on a Likert scale ranging from 0 ("cannot do") to 4 ("no difficulty"), the outcome was modeled by an order-logit response function. The ordered-logit is a generalization of the binary logistic response function that was designed to handle ordinal variables (Liddell and Kruschke, 2018; Bürkner and Vuorre, 2019). The results of an ordered-logit model consist of regression parameters for effect estimates on a logit scale (similar to common logistic regression).

The model was fitted using the Hamiltonian Monte Carlo (HMC) sampling algorithm in Stan version 2.21.0 (Stan Development Team, 2021) accessed *via* brms package (Bürkner and Vuorre, 2019) in R version 4.0.5 (R Core Team, 2021) using four independent chains with 1,500 total and 500 warm-up iterations. Full Bayesian statistical inference was used to specify the model and evaluate the results. We used Student-*t* priors with zero mean, a scale of 2.5 and three degrees of freedom for Intercepts and random effects' variance components and regularizing Normal priors with zero mean and standard deviation of 0.5 for the fixed effects. GLMM parameters were described on a logit scale by their medians, 95% highest density posterior probability intervals (PPIs) and the probability of being positive (i.e., the probability that a predictor has a positive effect on IADL). A 95% PPI can be interpreted such that a given parameter lies within this interval with a 95% probability. If desired, an effect can be regarded statistically significant (on a 5% level) if the corresponding 95% PPI excludes zero. To evaluate post-surgery change in IADL on the outcome scale we provide posterior predictions comparing the contrast between post-surgery minus pre-surgery responses to PDAQ across patients and items. Scripts with all analyses from this article are deposited here: https://github.com/josefmana/dbs_postop_iADL.

RESULTS

Characterizing the Sample

Characteristics of the sample are presented in **Table 1**. A total of 32 patients with PD and bilateral STN-DBS implanted with DBS devices between 2018 and 2019 met the inclusion criteria. Patients' responses to each PDAQ item before and after the STN-DBS surgery are depicted in **Supplementary Figure 1**.

Results of the Generalized Linear Mixed Model

The HMC sampling algorithm successfully converged to stable posterior distribution (all $\hat{R}$s < 1.01). Fixed effects' parameters are presented in **Table 2**. There was a 99.4% probability that patients report improved IADL post- as compared to pre-surgery when covariates are kept constant (the main effect of the time of assessment). Moreover, 95% PPI of this effect excluded zero and the effect can thus be regarded as statistically significant. Similarly, for DRS-2, there was a 98.8% probability

TABLE 2 | Fixed effect parameters of the ordered-logit generalized linear mixed model.

Predictor	*b*	95% PPI	Pr(*b* > 0)
Time of assessment	0.72	[0.21, 1.32]	0.994
LEDD	0.12	[−0.10, 0.35]	0.861
DRS-2	0.31	[0.02, 0.56]	0.988
BDI-II	−0.26	[−0.54, 0.01]	0.031
Time of assessment × LEDD	0.20	[−0.10, 0.51]	0.892
Time of assessment × DRS-2	−0.13	[−0.40, 0.17]	0.188
Time of assessment × BDI-II	−0.14	[−0.48, 0.17]	0.204

×, statistical interaction; b, median parameter estimate; BDI-II, Beck Depression Inventory; DRS-2, Dementia Rating Scale, second edition; LEDD, levodopa equivalent daily dose; PPI, highest density posterior probability interval; Pr(b > 0), probability that the parameter is positive (i.e., the effect "helps" with IADL evaluated by PDAQ, range 0–1); Time of assessment, pre- vs. post-surgery variable (higher values indicated post-surgery improvement). The time of assessment was deviation coded (i.e., pre-surgery = −0.5, post-surgery = 0.5) such that the main effects of DRS-2 and LEDD reflect the average effects across pre- and post-surgery assessments.

that it has a positive main effect on IADL (i.e., higher scores in DRS-2 positively affected IADL regardless of the time of assessment) with 95% PPI excluding zero. On the other hand, while both LEDD and BDI-II showed a trend of the positive main effect on IADL, their 95% PPIs included zero. There was a trend of an interaction between the time of assessment and LEDD, DRS-2, and BDI-II. However, the 95% PPI for these effects included zero as well as moderate negative values and these effects thus cannot be regarded as statistically significant.

Figure 2 depicts the main effects of assessment time, LEDD, DRS-2, and BDI-II on responses to PDAQ on the outcome scale (i.e., probabilities that a patient responds with each of the options 0–4). **Figure 2** shows that when evaluating difficulties in IADL patients rarely selected options 0–2 ("cannot do," "a lot," and "somewhat," respectively). The main effects of the time of assessment, DRS-2, BDI-II, and LEDD were primarily due to increased frequency of response four ("none") at the expense of response three ("a little") after the surgery and in patients with high LEDD and DRS-2 and low BDI-II scores. When LEDD, DRS-2, and BDI-II were statistically held at the average in-sample pre-surgery level, the post-surgery probability that a patient responds to any PDAQ item with option zero ("cannot do") decreased by 0.1% (95% PPI [−0.4, 0.0], pre: 0.3% [0.1, 0.7], post: 0.1% [0.0, 0.4]), the probability of response one ("a lot") decreased by 0.4% (95% PPI [−1.2, −0.1], pre: 0.9% [0.3, 2.1], post: 0.5% [0.1, 1.1]), the probability of response two ("somewhat") decreased by 1.9% (95% PPI [−4.6, −0.3], pre: 4.1% [1.7, 8.4], post: 2.1% [0.6, 4.7]), the probability of response three ("a little") decreased by 12.3% (95% PPI [−22.2, −3.4], pre: 36.1% [21.8, 51.1], post: 22.9% [10.1, 39.5]), and the probability of response four ("none") increased by 15.0% (95% PPI [3.1, 26.4], pre: 58.5% [39.2, 77.1], post: 74.4% [55.6, 90.1]). In other words, the direct effect of DBS on IADL improvement is due to a significantly lower frequency of patients with little IADL difficulties (response three "a little" in all PDAQ items) and a significantly higher frequency of patients with no IADL difficulties (response four "none" in all PDAQ items) 1 year after the STN-DBS surgery. This pattern of responses holds across different LEDD levels (see **Table 3**).

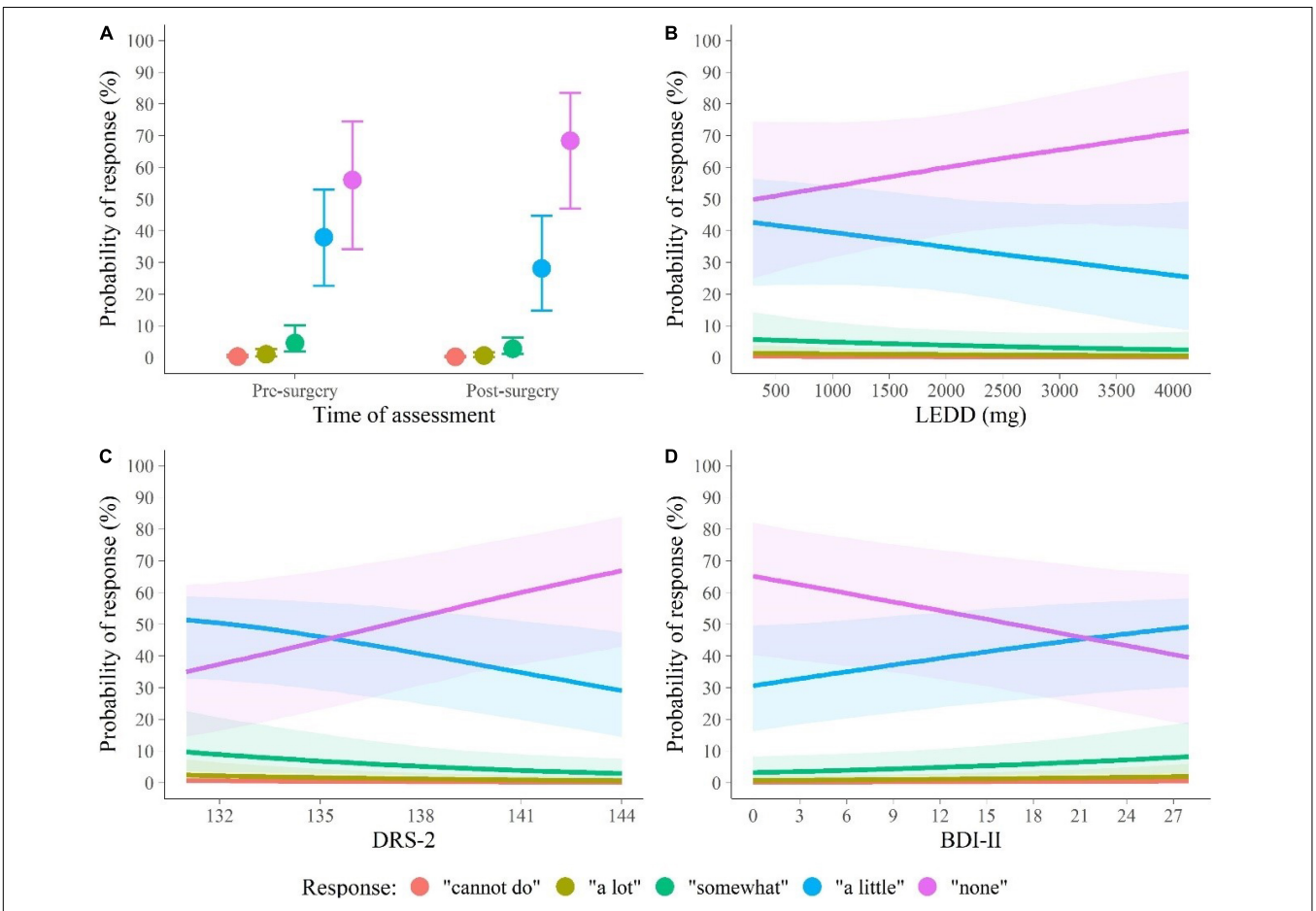


FIGURE 2 | Summary of the marginal distributions of main effects of **(A)** the time of assessment, **(B)** LEDD, **(C)** DRS-2, and **(D)** BDI-II on IADL. BDI-II, Beck Depression Inventory; DRS-2, Dementia Rating Scale, second edition; IADL, instrumental activities of daily living; LEDD, levodopa equivalent daily dose. Points (lines) represent the median probability of each response (labeled by distinct colors) to the items of Penn Parkinson's Daily Activities Questionnaire-15 (PDAQ-15), and whiskers (shades) represent 95% posterior probability intervals (PPIs).

DISCUSSION

This study examined IADL in a cohort of PD patients undergoing STN-DBS treatment by using pre-test (pre-surgery) and post-test (post-surgery) measurements. Regarding the complexity of DBS neurosurgical treatment, only IADL self-report and cognitive, depressive, and clinical correlates were included in the current research. In comparison to other studies that concentrated either on IADL predictors or the quantification of the degree of IADL deficits in PD-NC in comparison to PD-MCI (Rosenthal et al., 2010; Pirogovsky et al., 2013; Foster, 2014; Fellows and Schmitter-Edgecombe, 2019; Becker et al., 2020; Cholerton et al., 2020; Foster and Doty, 2021; Schmitter-Edgecombe et al., 2021; Becker et al., 2022), or comparison and development of specific methods and sensitive IADL items for PD (Brennan et al., 2016a,b; Fellows and Schmitter-Edgecombe, 2019; Sulzer et al., 2020; Schmitter-Edgecombe et al., 2021), whereas our study focused selectively on the comparison of IADL in PD before and after STN-DBS with a PD-specific questionnaire (Brennan et al., 2016a,b). We based our estimate of the post-surgery IADL change on an explicit causal model allowing for easier model criticism and derivation of proper covariates to include in our model (Pearl, 2009; McElreath, 2020).

Based on our model and data, the post-surgery IADL of PD patients improves compared to the pre-surgery level. At the same time, higher cognitive performance and higher LEDD are indicative of higher IADL both before and after STN-DBS surgery. More specifically, our analysis focused on the direct effects of all included predictors, in other words, the estimation of IADL post-surgery improvement was thus adjusted on LEDD, DRS-2, and BDI-II. While this approach allowed us to derive a more accurate estimate of the direct effect DBS can have on IADL performance in a cohort of implanted patients, this effect will be in the real-life clinical settings contaminated by DBS effects on other variables predictive of IADL change. Indeed, according to our results, the IADL declines when LEDD is decreased both before and after surgery, however, a decrease in LEDD is often a desirable outcome of DBS treatment (Molinuevo et al., 2000; Russmann et al., 2004). Surprisingly, there was a trend of an interaction between the time of assessment and LEDD indicating that higher LEDD may be more important for IADL improvement post-surgery than pre-surgery.

As a consequence of the above described putatively opposing effects of DBS and post-surgery LEDD reduction, medical professionals may want to carefully consider how much to reduce the LEDD after STN-DBS surgery in PD patients to avoid negative effects on IADL. In the current study, these considerations are quantitatively represented in **Table 3** which can be used to guide decisions on how much to decrease the LEDD after STN-DBS surgery while avoiding adverse effects

TABLE 3 | Expected response probabilities of difficulty in IADL stratified by the time of assessment and levodopa equivalent daily dose derived from the ordered-logit GLMM.

Assessment	LEDD (mg)	Pr(resp = 0)	Pr(resp = 1)	Pr(resp = 2)	Pr(resp = 3)	Pr(resp = 4)
Pre-surgery	0	0.5 ± 0.4%	1.7 ± 1.2%	7.0 ± 3.8%	43.1 ± 9.2%	47.6 ± 13.6%
	500	0.4 ± 0.3%	1.5 ± 0.9%	6.1 ± 3.0%	41.5 ± 8.6%	50.4 ± 12.1%
	1,000	0.4 ± 0.3%	1.3 ± 0.7%	5.4 ± 2.4%	39.6 ± 8.0%	53.3 ± 10.9%
	1,500	0.3 ± 0.2%	1.1 ± 0.6%	4.8 ± 2.0%	37.5 ± 7.7%	56.3 ± 10.1%
	2,000	0.3 ± 0.2%	1.0 ± 0.5%	4.3 ± 1.8%	35.3 ± 7.7%	59.1 ± 9.9%
	2,500	0.3 ± 0.2%	0.9 ± 0.5%	3.9 ± 1.8%	33.1 ± 8.1%	61.9 ± 10.2%
	3,000	0.2 ± 0.2%	0.8 ± 0.5%	3.5 ± 1.8%	31.0 ± 8.7%	64.4 ± 10.9%
	3,500	0.2 ± 0.2%	0.8 ± 0.5%	3.2 ± 1.9%	29.0 ± 9.5%	66.8 ± 11.8%
	4,000	0.2 ± 0.2%	0.7 ± 0.5%	3.0 ± 2.0%	27.2 ± 10.4%	68.9 ± 12.8%
	4,500	0.2 ± 0.2%	0.7 ± 0.6%	2.8 ± 2.2%	25.5 ± 11.2%	70.9 ± 13.8%
	5,000	0.2 ± 0.2%	0.6 ± 0.6%	2.7 ± 2.4%	23.9 ± 12.0%	72.6 ± 14.8%
Post-surgery	0	0.5 ± 0.3%	1.6 ± 0.9%	6.7 ± 3.0%	43.2 ± 7.9%	48.0 ± 11.5%
	500	0.3 ± 0.2%	1.2 ± 0.6%	5.0 ± 2.1%	38.1 ± 7.8%	55.5 ± 10.3%
	1,000	0.2 ± 0.2%	0.9 ± 0.4%	3.7 ± 1.6%	32.4 ± 7.7%	62.8 ± 9.6%
	1,500	0.2 ± 0.1%	0.6 ± 0.3%	2.8 ± 1.3%	27.0 ± 7.7%	69.4 ± 9.2%
	2,000	0.1 ± 0.1%	0.5 ± 0.3%	2.1 ± 1.1%	22.0 ± 7.8%	75.2 ± 9.2%
	2,500	0.1 ± 0.1%	0.4 ± 0.3%	1.6 ± 1.0%	17.9 ± 7.9%	80.1 ± 9.2%
	3,000	0.1 ± 0.1%	0.3 ± 0.2%	1.3 ± 1.0%	14.4 ± 7.9%	83.9 ± 9.1%
	3,500	0.1 ± 0.1%	0.2 ± 0.2%	1.0 ± 0.9%	11.7 ± 7.8%	87.0 ± 8.9%
	4,000	0.1 ± 0.1%	0.2 ± 0.2%	0.8 ± 0.9%	9.5 ± 7.6%	89.4 ± 8.7%
	4,500	0.0 ± 0.1%	0.2 ± 0.2%	0.7 ± 0.9%	7.8 ± 7.4%	91.3 ± 8.5%
	5,000	0.0 ± 0.1%	0.1 ± 0.3%	0.5 ± 0.9%	6.5 ± 7.1%	92.8 ± 8.2%

GLMM, generalized linear mixed model; IADL, instrumental activities of daily living; LEDD, levodopa equivalent daily dose; Pr(resp = i), probability that a patient will respond to any item of The Penn Parkinson's Daily Activities Questionnaire-15 (PDAQ-15) with the response "i" where "i" represents difficulties in IADL and can take on values 0 = "cannot do," 1 = "a lot," 2 = "somewhat," 3 = "a little," and 4 = "none"; the numbers represent posterior predictions of the ordered-logit GLMM for a patient with an average cognitive performance (Dementia Rating Scale, DRS-2≈139) and level of depressive symptoms (Beck Depression Inventory, BDI-II≈10) described in the main text in a format mean ± standard deviation.

on IADL. For instance, based on **Table 3**, one can expect that a patient with pre-surgery LEDD of 2,500 mg will report no difficulties (response four) in IADL about 62% of the time while reporting little difficulties (response three) about 33% of the time. If a physician was to reduce this patient's LEDD after the surgery to 1,000 mg, one can expect that patient's IADL would remain similar to the pre-surgery level reporting no difficulties about 63% of the time and little difficulties about 33% of the time. However, if the LEDD was discontinued altogether, the expectation of IADL difficulties would increase to a level where one would expect the patient to report no problems only 48% of the time and little problems about 43% of the time. In this case, it would thus be advisable not to reduce LEDD below 1,000 mg if the patient wanted to avoid possible adverse effects on IADL. This finding can be considered as an example of the "masked" effect (McElreath, 2020) with LEDD playing a crucial role in modulating the effect of STN-DBS on IADLs.

The present study suffers from several limitations that must be clearly stated. First, we did not apply multiple IADL assessment methods (i.e., observed everyday activities, self- and informant-reports) which would show different facets of IADL (Fellows and Schmitter-Edgecombe, 2019; Schmitter-Edgecombe et al., 2021). However, it is questionable in our PD sample with PD-NC or PD-MCI in the early stages of cognitive decline if self-rating is not more sensitive to the impact of cognitive changes on IADL function than informant reports (Cholerton et al., 2020). Second, our research is not longitudinal and we are not able to trace long-lasting changes in IADL due to STN-DBS. Third, we do not report data on individual stimulation volumes and functional zones of the STN and their contribution to IADL changes (Tödt et al., 2022). Fourth, a modest sample size regarding DBS research and the apparent lack of cognitive decline after STN-DBS surgery in our sample might have prevented us from observing any significant interaction between the time of assessment and cognitive performance. Fifth, an important limitation of the study is the lack of a control group. The influence of LEDD and global cognition on the improvement of IADL at 1 year of follow-up could be independent of the effect of DBS on motor symptoms. Such comparison cannot be performed due to a lack of a control group. However, based on previous research, the neurostimulation, as compared with medication alone, caused greater improvements from baseline to 6 months and DBS and levodopa have a differential effect on brain motor activity in PD (Deuschl et al., 2006; Mueller et al., 2020).

The current study shows a clear beneficial STN-DBS-induced change in IADL approximately 1-year perspective after the operation. Importantly, we show the IADL post-surgery improvement is related also to LEDD post-surgery medication dose that should not decrease under a certain limit to maintain the positive IADL effect of the surgery. Based on our study, STN-DBS seems as a cognitively safe procedure for the treatment of motor symptoms in PD 1 year after the surgery, however, a lower global cognitive functioning in the pre-surgery phase is associated with lower IADL functioning before and after the operation. Understanding the role of IADL functioning in PD in the pre-surgery phase may help identify those at risk for everyday activities and possibly help to improve interventions to promote functional independence after the electrode implantation.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the General University Hospital in Prague, Czechia. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

OB: conceptualization, data curation, investigation, methodology, supervision, writing—original draft and review and editing. JM: conceptualization, data curation, investigation, formal analysis, methodology, software, visualization, writing—original draft and review and editing. FR, FH, AF, TU, and DU: investigation and writing—review and editing. ER: conceptualization, funding acquisition, investigation, resources, and writing—review and editing. RJ: conceptualization, data acquisition and curation, funding acquisition, investigation, resources, supervision, and writing—review and editing. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Institute for Neurological Research, Czech Republic, Programme EXCELES, ID Project No. LX22NPO5107, funded by the European Union—Next Generation EU and also by the Charles University: Cooperatio Program in Neuroscience, also by AZV NV19-04-00233 offered by the Czech Ministry of Health, GA UK 254121 offered by Grant Agency of Charles University and JPND 733051123 offered by the EU Joint Programme on Neurodegenerative Disease Research.

ACKNOWLEDGMENTS

We would like to thank all the patients, family members, and staff from all the units that participated in the study. In particular, we would also like to thank Markéta Fialová, Radka Steinbachová, and Anna Rezková for their management of data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.886491/full#supplementary-material

REFERENCES

Altieri, M., Garramone, F., and Santangelo, G. (2021). Functional autonomy in dementia of the Alzheimer's type, mild cognitive impairment, and healthy aging: a meta-analysis. *Neurol. Sci.* 42, 1773–1783. doi: 10.1007/s10072-021-05142-0

Beck, A. T., Steer, R. A., and Brown, G. (1996). *Manual For The Beck Depression Inventory-II*. San Antonio, TX: Pearson.

Becker, S., Bäumer, A., Maetzler, W., Nussbaum, S., Timmers, M., Van Nueten, L., et al. (2020). Assessment of cognitive-driven activity of daily living impairment in non-demented Parkinson's patients. *J. Neuropsychol.* 14, 69–84. doi: 10.1111/jnp.12173

Becker, S., Pauly, C., Lawton, M., Hipp, G., Bowring, F., Sulzer, P., et al. (2022). Quantifying activities of daily living impairment in Parkinson's disease using the Functional Activities Questionnaire. *Neurol. Sci.* 43, 1047–1054. doi: 10.1007/s10072-021-05365-1

Bezdicek, O., Michalec, J., Nikolai, T., Havrankova, P., Roth, J., Jech, R., et al. (2015). Clinical validity of the Mattis Dementia Rating Scale in differentiating mild cognitive impairment in Parkinson's disease and normative data. *Dement. Geriatr. Cogn. Disord.* 39, 303–311. doi: 10.1159/000375365

Bezdicek, O., Nikolai, T., Michalec, J., Růžička, F., Havránková, P., Roth, J., et al. (2016). The diagnostic accuracy of parkinson's disease mild cognitive impairment battery using the movement disorder society task force criteria. *Mov. Disord. Clin. Pract.* 4, 237–244. doi: 10.1002/mdc3.12391

Bezdicek, O., Sulc, Z., Nikolai, T., Stepankova, H., Kopecek, M., Jech, R., et al. (2017). A parsimonious scoring and normative calculator for the Parkinson's disease mild cognitive impairment battery. *Clin. Neuropsychol.* 31, 1231–1247. doi: 10.1080/13854046.2017.1293161

Bratsos, S., Karponis, D., and Saleh, S. N. (2018). Efficacy and Safety of Deep Brain Stimulation in the Treatment of Parkinson's Disease: a systematic review and meta-analysis of randomized controlled trials. *Cureus* 10:e3474. doi: 10.7759/cureus.3474

Brennan, L., Siderowf, A., Rubright, J. D., Rick, J., Dahodwala, N., Duda, J. E., et al. (2016a). Development and initial testing of the penn Parkinson's daily activities questionnaire. *Mov Disord* 31, 126–134. doi: 10.1002/mds.26339

Brennan, L., Siderowf, A., Rubright, J. D., Rick, J., Dahodwala, N., Duda, J. E., et al. (2016b). The penn Parkinson's daily activities questionnaire-15: psychometric properties of a brief assessment of cognitive instrumental activities of daily living in Parkinson's disease. *Parkinsonism Relat. Disord.* 25, 21–26. doi: 10.1016/j.parkreldis.2016.02.020

Bronnick, K., Ehrt, U., Emre, M., De Deyn, P. P., Wesnes, K., Tekin, S., et al. (2006). Attentional deficits affect activities of daily living in dementia-associated with Parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 77, 1136–1142. doi: 10.1136/jnnp.2006.093146

Bronstein, J. M., Tagliati, M., Alterman, R. L., Lozano, A. M., Volkmann, J., Stefani, A., et al. (2011). Deep brain stimulation for Parkinson disease: an expert consensus and review of key issues. *Arch. Neurol.* 68:165. doi: 10.1001/archneurol.2010.260

Bürkner, P.-C., and Vuorre, M. (2019). Ordinal regression models in psychology: a tutorial. *Adv. Methods Pract. Psychol. Sci.* 2, 77–101.

Cahn, D. A., Sullivan, E. V., Shear, P. K., Pfefferbaum, A., Heit, G., and Silverberg, G. (1998). Differential contributions of cognitive and motor component processes to physical and instrumental activities of daily living in Parkinson's disease. *Arch. Clin. Neuropsychol.* 13, 575–583.

Cholerton, B., Poston, K. L., Tian, L., Quinn, J. F., Chung, K. A., Hiller, A. L., et al. (2020). Participant and study partner reported impact of cognition on functional activities in parkinson's disease. *Mov. Disord. Clin. Pract.* 7, 61–69. doi: 10.1002/mdc3.12870

Christ, J. B., Fruhmann Berger, M., Riedl, E., Prakash, D., Csoti, I., Molt, W., et al. (2013). How precise are activities of daily living scales for the diagnosis of Parkinson's disease dementia? A pilot study. *Parkinsonism Relat. Disord.* 19, 371–374. doi: 10.1016/j.parkreldis.2012.11.004

Ciharova, M., Cígler, H., Dostálová, V., Šivicová, G., and Bezdicek, O. (2020). Beck depression inventory, second edition, Czech version: demographic correlates, factor structure and comparison with foreign data. *Int. J. Psychiatry Clin. Pract.* 24, 371–379. doi: 10.1080/13651501.2020.1775854

Deuschl, G., Schade-Brittinger, C., Krack, P., Volkmann, J., Schäfer, H., Bötzel, K., et al. (2006). A randomized trial of deep-brain stimulation for Parkinson's disease. *N. Engl. J. Med.* 355, 896–908. doi: 10.1056/NEJMoa060281

Emre, M., Aarsland, D., Brown, R., Burn, D. J., Duyckaerts, C., Mizuno, Y., et al. (2007). Clinical diagnostic criteria for dementia associated with Parkinson's disease. *Mov. Disord.* 22, 1689–1707; quiz1837. doi: 10.1002/mds.21507

Fellows, R. P., and Schmitter-Edgecombe, M. (2019). Multimethod assessment of everyday functioning and memory abilities in Parkinson's disease. *Neuropsychology* 33, 169–177. doi: 10.1037/neu0000505

Foley, J. A., Foltynie, T., Limousin, P., and Cipolotti, L. (2018). Standardised neuropsychological assessment for the selection of patients undergoing DBS for Parkinson's Disease. *Parkinsons Dis.* 2018:4328371. doi: 10.1155/2018/4328371

Foster, E. R. (2014). Instrumental activities of daily living performance among people with Parkinson's disease without dementia. *Am. J. Occup. Ther.* 68, 353–362. doi: 10.5014/ajot.2014.010330

Foster, E. R., and Doty, T. (2021). Cognitive correlates of instrumental activities of daily living performance in parkinson disease without dementia. *Arch. Rehabil. Res. Clin. Transl.* 3:100138. doi: 10.1016/j.arrct.2021.100138

Giovannetti, T., Britnell, P., Brennan, L., Siderowf, A., Grossman, M., Libon, D. J., et al. (2012). Everyday action impairment in Parkinson's disease dementia. *J. Int. Neuropsychol. Soc.* 18, 787–798. doi: 10.1017/S135561771200046X

Gorecka-Mazur, A., Furgala, A., Krygowska-Wajs, A., Pietraszko, W., Kwinta, B., and Gil, K. (2019). Activities of daily living and their relationship to health-related quality of life in patients with parkinson disease after subthalamic nucleus deep brain stimulation. *World Neurosurg.* 125, e552–e562. doi: 10.1016/j.wneu.2019.01.132

Hariz, G.-M., and Forsgren, L. (2011). Activities of daily living and quality of life in persons with newly diagnosed Parkinson's disease according to subtype of disease, and in comparison to healthy controls. *Acta Neurol. Scand.* 123, 20–27. doi: 10.1111/j.1600-0404.2010.01344.x

Hentz, J. G., Mehta, S. H., Shill, H. A., Driver-Dunckley, E., Beach, T. G., and Adler, C. H. (2015). Simplified conversion method for unified Parkinson's disease rating scale motor examinations. *Mov. Disord.* 30, 1967–1970. doi: 10.1002/mds.26435

Hoogland, J., Boel, J. A., De Bie, R. M. A., Geskus, R. B., Schmand, B. A., Dalrymple-Alford, J. C., et al. (2017). Mild cognitive impairment as a risk factor for Parkinson's disease dementia. *Mov. Disord.* 32, 1056–1065. doi: 10.1002/mds.27002

Hughes, A. J., Daniel, S. E., Kilford, L., and Lees, A. J. (1992). Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J. Neurol. Neurosurg. Psychiatry* 55, 181–184. doi: 10.1136/jnnp.55.3.181

Jech, R., Mueller, K., Urgošík, D., Sieger, T., Holiga, Š, Růžička, F., et al. (2012). The Subthalamic microlesion story in Parkinson's disease: electrode insertion-related motor improvement with relative cortico-subcortical hypoactivation in fMRI. *PLoS One* 7:e49056. doi: 10.1371/journal.pone.0049056

Jech, R., Ruzicka, E., Urgosik, D., Serranova, T., Volfova, M., Novakova, O., et al. (2006). Deep brain stimulation of the subthalamic nucleus affects resting EEG and visual evoked potentials in Parkinson's disease. *Clin. Neurophysiol.* 117, 1017–1028. doi: 10.1016/j.clinph.2006.01.009

Jiang, J.-L., Chen, S.-Y., Hsieh, T.-C., Lee, C.-W., Lin, S.-H., and Tsai, S.-T. (2015). Different effectiveness of subthalamic deep brain stimulation in Parkinson's disease: a comparative cohort study at 1 year and 5 years. *J. Formos. Med. Assoc.* 114, 835–841. doi: 10.1016/j.jfma.2013.09.006

Jurica, P. J., Leitten, C. L., and Mattis, S. (2001). *Dementia Rating Scale-2 (DRS-2) Professional Manual*. Lutz, FL: Psychological Assessment Resources.

Kleiner-Fisman, G., Stern, M. B., and Fisman, D. N. (2010). Health-related quality of life in parkinson disease: correlation between health utilities index iii and unified parkinson's disease rating scale (UPDRS) in U.S. male veterans. *Health Qual. Life Outcomes* 8:91. doi: 10.1186/1477-7525-8-91

Kubu, C. S. (2018). The role of a neuropsychologist on a movement disorders deep brain stimulation team. *Arch. Clin. Neuropsychol.* 33, 365–374. doi: 10.1093/arclin/acx130

Lawson, R. A., Williams-Gray, C. H., Camacho, M., Duncan, G. W., Khoo, T. K., Breen, D. P., et al. (2021). Which neuropsychological tests? Predicting cognitive decline and dementia in Parkinson's Disease in the ICICLE-PD cohort. *J. Parkinsons Dis.* 11, 1297–1308. doi: 10.3233/JPD-212581

Liddell, T. M., and Kruschke, J. K. (2018). Analyzing ordinal data with metric models: what could possibly go wrong? *J. Exp. Soc. Psychol.* 79, 328–348.

Litvan, I., Goldman, J. G., Tröster, A. I., Schmand, B. A., Weintraub, D., Petersen, R. C., et al. (2012). Diagnostic criteria for mild cognitive impairment in Parkinson's disease: movement disorder society task force guidelines. *Mov. Disord.* 27, 349–356. doi: 10.1002/mds.24893

Martin, R. C., Triebel, K. L., Kennedy, R. E., Nicholas, A. P., Watts, R. L., Stover, N. P., et al. (2013). Impaired financial abilities in Parkinson's disease patients with mild cognitive impairment and dementia. *Parkinsonism Relat. Disord.* 19, 986–990. doi: 10.1016/j.parkreldis.2013.06.017

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course With Examples in R and Stan*. London: Chapman and Hall/CRC.

Molinuevo, J. L., Valldeoriola, F., Tolosa, E., Rumià, J., Valls-Solé, J., Roldán, H., et al. (2000). Levodopa withdrawal after bilateral subthalamic nucleus stimulation in advanced parkinson disease. *Arch. Neurol.* 57, 983–988. doi: 10.1001/archneur.57.7.983

Moro, E., and Lang, A. E. (2006). Criteria for deep-brain stimulation in Parkinson's disease: review and analysis. *Exp. Rev. Neurother.* 6, 1695–1705. doi: 10.1586/14737175.6.11.1695

Mueller, K., Urgošík, D., Ballarini, T., Holiga, Š, Möller, H. E., Růžička, F., et al. (2020). Differential effects of deep brain stimulation and levodopa on brain activity in Parkinson's disease. *Brain Commun.* 2:fcaa005. doi: 10.1093/braincomms/fcaa005

Odekerken, V. J. J., Van Laar, T., Staal, M. J., Mosch, A., Hoffmann, C. F. E., Nijssen, P. C. G., et al. (2013). *Subthalamic nucleus* versus *Globus pallidus* bilateral deep brain stimulation for advanced Parkinson's disease (NSTAPS study): a randomised controlled trial. *Lancet Neurol.* 12, 37–44. doi: 10.1016/S1474-4422(12)70264-8

Okun, M. S. (2014). Deep-brain stimulation–entering the era of human neural-network modulation. *N. Engl. J. Med.* 371, 1369–1373. doi: 10.1056/NEJMp1408779

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference.* Cambridge: Cambridge University Press.

Perlmutter, J. S., and Mink, J. W. (2006). Deep brain stimulation. *Ann. Rev. Neurosci.* 29, 229–257.

Pirogovsky, E., Martinez-Hannon, M., Schiehser, D. M., Lessig, S. L., Song, D. D., Litvan, I., et al. (2013). Predictors of performance-based measures of instrumental activities of daily living in nondemented patients with Parkinson's disease. *J. Clin. Exp. Neuropsychol.* 35, 926–933. doi: 10.1080/13803395.2013.838940

Pirogovsky, E., Schiehser, D. M., Obtera, K. M., Burke, M. M., Lessig, S. L., Song, D. D., et al. (2014). Instrumental activities of daily living are impaired in Parkinson's disease patients with mild cognitive impairment. *Neuropsychology* 28, 229–237. doi: 10.1037/neu0000045

R Core Team (2021). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing.

Rosenthal, E., Brennan, L., Xie, S., Hurtig, H., Milber, J., Weintraub, D., et al. (2010). Association between cognition and function in patients with Parkinson disease with and without dementia. *Mov. Disord.* 25, 1170–1176. doi: 10.1002/mds.23073

Russmann, H., Ghika, J., Combrement, P., Villemure, J. G., Bogousslavsky, J., Burkhard, P. R., et al. (2004). L-Dopa-induced dyskinesia improvement after STN-DBS depends upon medication reduction. *Neurology* 63, 153–155. doi: 10.1212/01.wnl.0000131910.72829.9d

Schmitter-Edgecombe, M., McAlister, C., and Greeley, D. (2021). A comparison of functional abilities in individuals with mild cognitive impairment and Parkinson's disease with mild cognitive impairment using multiple assessment methods. *J. Int. Neuropsychol. Soc.* 1–12. doi: 10.1017/S1355617721001077 [Epub ahead of print].

Shulman, L. M., Armstrong, M., Ellis, T., Gruber-Baldini, A., Horak, F., Nieuwboer, A., et al. (2016). Disability rating scales in parkinson's disease: critique and recommendations. *Mov. Disord.* 31, 1455–1465. doi: 10.1002/mds.26649

Soh, S.-E., Mcginley, J. L., Watts, J. J., Iansek, R., Murphy, A. T., Menz, H. B., et al. (2013). Determinants of health-related quality of life in people with Parkinson's disease: a path analysis. *Qual. Life Res.* 22, 1543–1553. doi: 10.1007/s11136-012-0289-1

Stan Development Team (2021). *Stan Modeling Language User's Guide and Reference Manual (Version 2.21.0)*. Available online at: http://mc-stan.org/

Sulzer, P., Liebig, L., Csoti, I., Graessel, E., Wurster, I., Berg, D., et al. (2020). A time-efficient screening tool for activities of daily living functions in Parkinson's disease dementia. *J. Clin. Exp. Neuropsychol.* 42, 867–879. doi: 10.1080/13803395.2020.1825634

Tödt, I., Al-Fatly, B., Granert, O., Kühn, A. A., Krack, P., Rau, J., et al. (2022). The contribution of subthalamic nucleus deep brain stimulation to the improvement in motor functions and quality of life. *Mov. Disord.* 37, 291–301. doi: 10.1002/mds.28952

Tomlinson, C. L., Stowe, R., Patel, S., Rick, C., Gray, R., and Clarke, C. E. (2010). Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov. Disord.* 25, 2649–2653. doi: 10.1002/mds.23429

Urgosik, D., Jech, R., Ruzicka, E., and Ruzicka, F. (2011). Deep brain stimulation in movement disorders: a prague-center experience. *Cas Lek Cesk* 150, 223–228.

Young, T. L., Granic, A., Yu Chen, T., Haley, C. B., and Edwards, J. D. (2010). Everyday reasoning abilities in persons with Parkinson's disease. *Mov. Disord.* 25, 2756–2761. doi: 10.1002/mds.23379

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 22 June 2022
doi: 10.3389/fnagi.2022.833702

Enlarged Perivascular Spaces in the Basal Ganglia Independently Related to Gait Disturbances in Older People With Cerebral Small Vessel Diseases

Shuna Yang, Xuanting Li, Wenli Hu, Wei Qin and Lei Yang*

Department of Neurology, Beijing Chaoyang Hospital, Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:
Ondrej Bezdicek,
Charles University, Czechia

Reviewed by:
Helmet Karim,
University of Pittsburgh, United States
Qi Li,
The First Affiliated Hospital of Chongqing Medical University, China

****Correspondence:***
Wenli Hu
wenlihu3366@126.com

Specialty section:
This article was submitted to Parkinson's Disease and Aging-related Movement Disorders, a section of the journal Frontiers in Aging Neuroscience

Received: *12 December 2021*
Accepted: *31 May 2022*
Published: *22 June 2022*

Citation:
Yang S, Li X, Hu W, Qin W and Yang L (2022) Enlarged Perivascular Spaces in the Basal Ganglia Independently Related to Gait Disturbances in Older People With Cerebral Small Vessel Diseases. Front. Aging Neurosci. 14:833702. doi: 10.3389/fnagi.2022.833702

Background and Objective: Gait disturbances are common in older people and are associated with adverse consequences, e.g., falls and institutionalization. Enlarged perivascular spaces in the basal ganglia (BG-EPVS) are considered an magnetic resonance imaging (MRI) marker of cerebral small vessel diseases (CSVD). However, the consequences of BG-EPVS are largely unknown. Previous studies showed that other CSVD markers were related to gait disturbances. However, the relation between BG-EPVS and gait performance is unclear. Therefore, we aimed to explore the relation between BG-EPVS and gait performance in elderly individuals.

Methods: We recruited older people with CSVD in the Neurology Department of our hospital from December 1, 2020 to October 31, 2021. Participants with BG-EPVS > 20 on the unilateral side of the basal ganglia slice containing the maximum number were classified into the BG-EPVS group (n = 78), and the rest were classified into the control group (n = 164). Quantitative gait parameters and gait variability were provided by the Intelligent Device for Energy Expenditure and Activity (IDEEA; MiniSun, United States) gait analysis system. Semiquantitative gait assessment was measured with the Tinetti test. Point-biserial correlation and multivariate linear regression analysis were performed to investigate the association between BG-EPVS and gait performance.

Results: The BG-EPVS group had a slower gait speed and cadence, shorter stride length, longer stance phase percentage, smaller pre-swing angle and footfall, and lower Tinetti gait test and balance test scores compared with those in the control group ($P < 0.05$). There were no statistical differences in stride length variability and stride time variability between the two groups ($P > 0.05$). A correlation analysis showed that BG-EPVS were negatively related to gait speed, cadence, stride length, pre-swing angle, and footfall ($\gamma_{range} = -0.497$ to -0.237, $P < 0.001$) and positively related to stance phase percentage ($\gamma = 0.269$, $P < 0.001$). BG-EPVS was negatively related to the score of the Tinetti gait test ($\gamma = -0.449$, $P < 0.001$) and the balance test ($\gamma = -0.489$, $P < 0.001$). The multiple linear regression analysis indicated that BG-EPVS

was an independent risk factor for gait disturbances and poor balance after adjusting for confounders, including other CSVD markers.

Conclusion: Large numbers of BG-EPVS were independently related to gait disturbances in older people with CSVD. This finding provides information about the consequences of BG-EPVS and risk factors for gait disturbances.

Keywords: cerebral small vessel diseases, enlarged perivascular spaces, Virchow–Robin spaces, gait, elderly individuals

INTRODUCTION

Gait disturbances are a major issue in older people because they are prevalent and related to adverse consequences, such as falls, institutionalization, and death (Verghese et al., 2006; Abellan van Kan et al., 2009). There is an important clinical significance in exploring the risk factors for and mechanisms underlying gait disturbances in older people. Normal gait control is a complex function that depends on the coordination of multiple brain regions, including cortical, subcortical, and spinal hubs (de Laat et al., 2012; Kim et al., 2016). The basal ganglia refers to a group of important subcortical strcutures related to the control of normal gait, balance, and falls. The motor circuit of the basal ganglia is important for the selection and suppression of movement, the execution of automatic actions, and the scaling of motor outputs (MacKinnon, 2018). Movement initiation and anticipatory adjustments of ongoing motion are achieved by cortical gait control. Automatic gait control is achieved by projections downstream of the mesencephalic locomotor region (MLR) and the pedunculopontine nucleus (PPN) to spinal central pattern generators that produce and modulate basic bipedal locomotor patterns. Tracts descending from the cortical areas project *via* the basal ganglia loop to adjust the activity of MLR/PPN through GABA-ergic inhibitory outputs of the pars reticulata of the substantia nigra (Wagner et al., 2022).

Perivascular spaces, or Virchow–Robin spaces, are compartments surrounding the small cerebral penetrating vessels, serving as a protolymphatic system that plays an important role in interstitial fluid and solute clearance in the brain (Gouveia-Freitas and Bastos-Leite, 2021). Perivascular spaces dilate with the accumulation of interstitial fluids. Enlarged perivascular spaces, visible with magnetic resonance imaging (MRI), appear as punctate or linear signal intensities similar to cerebrospinal fluid (CSF) on all MRI sequences (Rudie et al., 2018). Enlarged perivascular spaces in the basal ganglia (BG-EPVS) are recognized as an MRI marker of cerebral small vessel diseases (CSVD) (Wardlaw et al., 2013). However, the consequences of BG-EPVS are largely unknown. Although they are generally considered clinically silent, some studies explored the association between EPVS and cognitive disturbances (Zhu et al., 2010; Jie et al., 2020). However, the results were not consistent across the studies. Whether the association between EPVS and cognitive impairment is attributed to a direct effect of EPVS or to the accompanying CSVD markers, such as white matter hyperintensities (WMH), cerebral microbleeds (CMB), and lacunes, remains incompletely understood.

Previous studies showed that WMH, CMB, and lacunes were related to gait disturbances (Stijntjes et al., 2016; van der Holst et al., 2018; Cannistraro et al., 2019). However, the relation between BG-EPVS and gait performance is unclear. We speculated that the presence of many BG-EPVS might be associated with gait disturbances by disrupting the function of the basal ganglia. Therefore, we explored the association between BG-EPVS and gait performance after adjusting for WMH, CMB, and lacunes in older individuals.

MATERIALS AND METHODS

Study Subjects

The study was a cross-sectional study. We reviewed the medical records of older patients with CSVD and without acute cerebral infarction who were admitted to the Neurology Department of Beijing Chaoyang Hospital affiliated with Capital Medical University from December 1, 2020 to October 31, 2021. CSVD was defined as the presence of WMH, CMB, and/or lacunes of presumed vascular origin on brain MRI. Patients were selected for participating in the study according to the following inclusion and exclusion criteria. The inclusion criteria were the following: (1) aged 60 or over and (2) agreed to participate in the study. The exclusion criteria were the following: (1) dementia including Alzheimer's disease, frontotemporal dementia, or dementia with Lewy bodies; (2) Parkinson's diseases (PD) and Parkinson's plus syndromes, including multiple system atrophy, cortical basal degeneration, and progressive supranuclear palsy; (3) history of severe stroke (the largest diameter of lesion size $>$ 20 mm) that caused difficulties and inaccurate assessments of CSVD MRI markers or large-vessel cerebrovascular diseases defined as internal carotid, middle cerebral, or basilar intracranial artery stenosis $>$ 50%; (4) lacunar syndrome within 6 months after the event to avoid acute effects on the outcomes, or the patients with stroke sequelae 6 months after cerebral infarction onset; (5) traumatic, toxic or infectious brain injury, brain tumor or brain metastases, or non-CSVD-rated white matter lesions, e.g., multiple sclerosis and irradiation induced gliosis; (6) inability to walk for 30 m unaided; (7) could not finish the tests because of prominent visual, hearing, or language impairments or psychiatric disease; (8) conditions not related to CSVD that affected gait (e.g., joint fusion, severe arthritis, joint replacement, or lumbar spondylopathy); (9) heart failure, myocardial infarction, or angina pectoris disorders during the previous 3 months or severe nephrosis or liver disease with a

life expectancy of < 6 months; and (10) MRI contraindications, known claustrophobia, or the patient's head moved during the MRI, resulting in poor imaging quality that affected the CSVD assessment.

All enrolled patients were divided into two groups according to the number of BG-EPVS on axial T2-weighted images. Patients with BG-EPVS > 20 on the unilateral side of the basal ganglia slice containing the maximum number were classified into the BG-EPVS group (**Figure 1A**); otherwise, the patients were classified into the control group (**Figure 1B**). The cutoff > 20 EPVS was used as a high grade, as in previous studies (Doubal et al., 2010).

Ethical Standards Statement

The study was approved by the Ethics Committee of Beijing Chaoyang Hospital Affiliated with Capital Medical University and was conducted in accordance with the Declaration of Helsinki. All participants provided informed written consent.

Demographic and Clinical Assessments

The following data were collected: age, sex, body mass index (BMI), past medical history, including hypertension, diabetes mellitus, hyperlipidemia, and stroke, and smoking and alcohol consumption. All blood samples were collected in the morning after an overnight fast and sent to the clinical laboratory of our hospital for the measurement of serum indices, including total cholesterol (TC), triglyceride, high-density lipoprotein (HDL), low-density lipoprotein (LDL), hemoglobin A1c (HbA1c), blood urea nitrogen (BUN), and creatinine.

Magnetic Resonance Imaging Examinations and Assessments of BG-EPVS, White Matter Hyperintensities, Cerebral Microbleeds, and Lacunes

Using a 3 T MRI scanner (Prisma; Siemens AG, Erlangen, Germany), MRI was performed in the Radiology Department of our hospital. The standardized MRI sequences included T1-weighted, T2-weighted, diffusion-weighted imaging, fluid-attenuated inversion recovery, and susceptibility-weighted imaging.

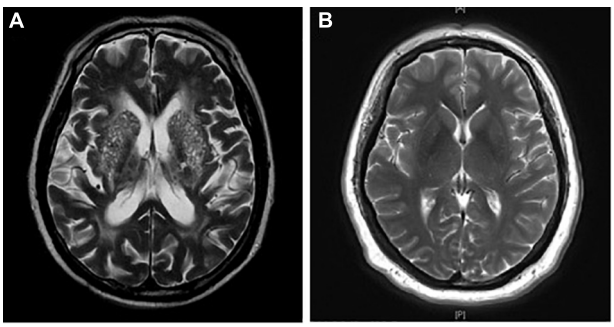


FIGURE 1 | Enlarged perivascular spaces in basal ganglia (BG-EPVS) group and control group. **(A)** Patients with EPVS > 20 on the unilateral side of the basal ganglia slice containing the maximum number were defined as the BG-EPVS group. **(B)** Patients with EPVS ≤ 20 on the unilateral side of the basal ganglia slice containing the maximum number were defined as the control group.

Imaging markers of CSVD, including BG-EPVS, WMH, CMB, and lacunes, were defined according to the Standards for Reporting Vascular Changes on Neuroimaging criteria described previously (Wardlaw et al., 2013). The number of BG-EPVS was assessed in the basal ganglia slice containing the maximum number of EPVS. Periventricular and deep WMH were evaluated separately and summed to obtain Fazekas scores. A detailed description of the Fazekas scale has been previously published (Fazekas et al., 1987). The numbers of CMB and lacune were counted.

Assessments of BG-EPVS, WMH, CMB, and lacunes were performed by two experienced neurologists blinded to the clinical data. Random scans of 50 individuals were independently examined by the two experienced neurologists blinded to the readings of the other. The intra-rater agreement for the ratings of BG-EPVS, WMH, CMB, and lacunes was assessed on a random sample of 50 individuals, with a 1-month interval between the first and second readings. The k statistics of the intra-rater and interrater agreements were 0.80 and 0.98, indicating good reliability. Disagreement was resolved by discussing it with the other co-authors.

Gait Measurement

Quantitative gait analysis was performed with the Intelligent Device for Energy Expenditure and Activity (IDEEA; MiniSun, United States), which had excellent test-retest reliability and validity (Gorelick et al., 2009; de la Cámara et al., 2019). The IDEEA was equipped with Windows-based software that analyzed the gait data and provided gait parameters including gait speed, cadence (number of steps per minute), stride length (the distance between the heel points of two consecutive footprints of the same foot), stance phase percentage, footfall (neuromuscular and skeletal control of limbs during the end of swing phase), and pre-swing angle (the foot-ground angle at the moment of toe-off). Gait variability included stride length variability and stride time variability, which reflected the magnitude of stride-to-stride fluctuations. Gait variability was calculated as the coefficient of variation as a percentage [(standard deviation of parameter/mean of parameter) × 100%]. The participants were instructed to walk 8 m wearing the device at their usual gait speed on low-heeled shoes before the formal test. The formal test of walking 8 m was then performed.

Semiquantitative gait assessment was measured using the Tinetti test with 17 items: nine for balance (score 0–16) and eight for gait (score 0–12) with a maximum score of 28 (de Laat et al., 2010). The Tinetti test was independently scored by two experienced neurologists blinded to the clinical data when the participants walked the 8 m; the two scores were averaged. Interrater reliability was calculated. The k coefficient of the interrater agreement was 0.96, indicating good reliability.

Statistical Analysis

Continuous variables were summarized as mean values ± *SD* or median (interquartile range) according to whether the data

conformed to a normal distribution. Categorical variables were presented as absolute numbers and percentages. Continuous variables with both normal distributions and homogeneity of variance were compared using the Student's *t*-test, whereas were compared using the Wilcoxon rank-sum test. Bonferroni correction was adopted to reset the *p*-value. The Chi-squared test was used for the comparison of categorical variables. Point-biserial correlation analysis was adopted to analyze the correlation between BG-EPVS (a binary variable) and gait parameters (continuous variables). The relationship between BG-EPVS and gait parameters was assessed with the point-biserial correlation coefficient. Partial correlation analysis and multivariate linear regression analysis were performed to analyze whether BG-EPVS was independently related to gait performance after adjusting for confounding factors. Analysis was performed with statistical product and service solutions 21.0 (SPSS 21.0, IBM, United States), and statistical significance was accepted at $P < 0.05$.

RESULTS

Participants' General Clinical Characteristics

A total of 354 elderly patients with CSVD and without acute cerebral infarction during the study period were identified. Among them, we excluded 43 patients who had a history of severe stroke and large-vessel diseases, 5 patients who had been diagnosed with PD, 36 patients who had severe arthritis, joint fusion, lumbar spondylopathy, or visual impairment that affected their walking, and 3 patients whose imaging quality was not high enough to assess markers of CSVD. The other 25 patients did not agree to participate in the study. Ultimately, 242 elderly patients were enrolled in the study and performed the tests, among which 78 patients were classified into the BG-EPVS group and 164 patients were classified into the control group according to the number of BG-EPVS.

The general clinical characteristics of all participants, the BG-EPVS group, and the control group are presented in **Table 1**. The mean age of all participants was 71 ± 7.7 years old and 141 (58.26%) of them were men. Those in the BG-EPVS group were older than those in the control group. The proportions of those of the male sex, with hypertension, and a history of stroke in the BG-EPVS group were higher than those in the control group. Those in the BG-EPVS group had higher levels of blood creatinine. There were no statistical differences between the BG-EPVS and control groups in BMI and the proportions of the following: smoking and alcohol, history of diabetes and hyperlipidemia, and the other collected clinical characteristics. Considering the imaging characteristics, the BG-EPVS group had more serious WMH, more lacunes, and CMB ($P < 0.001$). The detailed statistical results are shown in **Table 1**.

Association Between BG-EPVS and Gait Performance

Table 2 shows the gait parameters and statistical results of the BG-EPVS and control groups. Compared with the control group, the BG-EPVS group had a slower gait speed and cadence, shorter stride length, longer stance phase percentage, smaller pre-swing angle, and footfall. The scores of the Tinetti gait test and balance test were lower in BG-EPVS than those in the control group. Statistical differences were retained after Bonferroni correction.

Point-biserial correlation analysis showed that BG-EPVS were negatively related to gait speed, cadence, stride length, pre-swing angle, and foot control and positively related to stance phase percentage (**Table 3**). BG-EPVS was negatively related to the scores of the Tinetti gait test and the balance test. We further conducted a partial correlation analysis to adjust for the confounding factors. The negative correlation between BG-EPVS and gait performance and balance did not change even after adjusting for the confounders (**Table 3**).

To identify the effect of BG-EPVS on gait, we took the gait parameters and Tinetti test scores as dependent variables and conducted the multiple linear regression analysis. The results of the multiple linear regression analysis showed that BG-EPVS was an independent risk factor for gait disturbances and balance after adjusting for age, the proportion of male sex, hypertension, and stroke, level of blood creatinine, Fazekas score, the number of lacuna and CMB. The detailed analysis results are presented in **Table 4**.

Association Between BG-EPVS and Gait Variability

There were no statistical differences in stride length variability [27.6 (13–30.4) vs. 24.3 (11.7–27.7), $P = 0.08$] and stride time variability [11 (8.6–26.4) vs. 12.3 (8.9–30), $P = 0.374$] between the BG-EPVS group and the control group. Point-biserial correlation analysis also showed that BG-EPVS was not related to gait variability ($P > 0.05$).

DISCUSSION

In the present study, we explored the association between BG-EPVS and gait performance in older individuals with CSVD after adjusting for WMH, CMB, and lacunes. We found that large numbers of BG-EPVS were independently related to gait disturbances, including lower gait speed, shorter stride length, longer stance phase, smaller pre-swing angle, and poorer foot control using a quantitative gait analysis system. In addition, we found that BG-EPVS was also independently related to poor body balance. However, BG-EPVS was not related to gait variability.

The present clinical experience and previous studies indicated that, in subjects with CSVD, symptoms mainly arose in only moderate or severe cases (Baezner et al., 2008; de Laat et al., 2010). For example, gait disturbances were only attributed to moderate or severe WMH and/or the presence of > 3 lacunar infarcts (de Laat et al., 2010). Therefore, in the present study, we classified the subjects with high-grade BG-EPVS into the BG-EPVS group and others were classified into the control group.

Previous studies investigated the relationship between gait performance and other CSVD MRI markers, such as WMH, lacunar infarction, and CMB (Litak et al., 2020). The Radboud University Nijmegen Diffusion tensor and MRI Cohort (RUN DMC) study found that WMH and lacunar infarcts were both

independently associated with most gait parameters, and stride length was the most sensitive parameter related to WMH (de Laat et al., 2010). WMH in the basal ganglia and lacunar infarcts in the thalamus were related to a lower velocity. The RUN DMC study also found that a larger number of CMB in the basal ganglia was independently related to a shorter stride length and poorer performance on the Tinetti tests (de Laat et al., 2011b). However, the clinical studies specifically addressing the association between EPVS and gait in patients with CSVD were scarce. Shin et al. (2021) explored the relation between BG-EPVS and EPVS in white matter and motor function in patients with PD. They found that severe BG-EPVS was associated with worse motor symptoms. Chung et al. (2021) investigated the association between BG-EPVS and long-term motor outcomes in patients with PD. They found that the expression of the dopamine transporter was much lower, and the risk of freezing of gait was higher in the PD-EPVS + group. The PD-EPVS + group required higher doses of dopaminergic medications for effective symptom control than the PD-EPVS- group. In the present study, we explored the relationship between BG-EPVS and gait in older patients with CSVD. We found large numbers of BG-EPVS in older patients with CSVD independently related to worse gait performance, similar to the findings in previous studies. Su et al. (2017) investigated the

TABLE 1 | General clinical characteristics of participants.

	All participants	BG-EPVS group	Control group	T or Z or Chi-square value	*P*-value
n	242	78	164	–	–
Age[a], years	71 ± 7.7	75 ± 8.0	70 ± 7.0	-5.305	< 0.001
Men[c], *n* (%)	141 (58.26)	57 (73.08)	84 (51.22)	10.385	0.001
BMI[a], kg/m^2	25 ± 3.7	25 ± 3.5	25 ± 3.8	-0.065	0.948
Smoking[c], *n* (%)	77 (31.82)	30 (38.46)	47 (28.66)	2.342	0.126
Alcohol[c], *n* (%)	43 (17.77)	16 (20.51)	27 (16.46)	0.593	0.441
Hypertension[c], *n* (%)	170 (70.25)	62 (79.49)	108 (65.85)	4.701	0.030
Diabetes[c], *n* (%)	65 (26.86)	27 (34.62)	38 (23.17)	3.524	0.060
Hyperlipidemia[c], *n* (%)	89 (36.78)	29 (37.18)	60 (36.59)	0.008	0.929
Stroke[c], *n* (%)	54 (22.31)	30 (38.46)	24 (14.63)	17.312	< 0.001
TC[a], mmol/L	4.4 ± 0.99	4.3 ± 1.00	4.5 ± 0.99	1.091	0.276
Triglyceride[a], mmol/L	1.6 ± 0.92	1.6 ± 0.94	1.6 ± 0.90	0.088	0.930
HDL[a], mmol/L	1.1 ± 0.30	1.1 ± 0.29	1.1 ± 0.31	1.036	0.301
LDL[a], mmol/L	2.8 ± 0.96	2.7 ± 0.98	2.8 ± 0.94	0.907	0.365
HbA1[a], %	6.3 ± 1.27	6.6 ± 1.37	6.2 ± 1.21	-1.973	0.051
BUN[a], mmol/L	5.7 ± 1.72	6.0 ± 1.78	5.6 ± 1.68	-1.705	0.090
Creatinine[a], μmol/L	70 ± 23.2	75 ± 25.3	67 ± 21.6	-2.412	0.017
WMH[b], fazekas score	3 (2–4)	4 (3–5)	2 (2–3)	-8.195	< 0.001
Lacune[b], *n*	0 (0–2)	2 (0–4)	0 (0–1)	-7.546	< 0.001
CMB[b], *n*	0 (0–1)	1 (1–5)	0 (0–1)	-4.705	< 0.001

BMI, body mass index; TC, total cholesterol; HDL, high-density lipoprotein; LDL, low-density lipoprotein; HbA1c, hemoglobin A1c; BUN, blood urea nitrogen; WMH, white matter hyperintensities. [a]Continuous variables with normal distribution were expressed as mean values ± standard deviation and were compared with the Student t-test. The statistic value was T. [b]Continuous variables with non-normally distributions were expressed as median (interquartile range) and compared with the Wilcoxon rank-sum test. The statistic value was Z. [c]Categorical variables were presented as absolute numbers and percentages and compared with the Chi-squared test. The statistic value was the Chi-square value.

TABLE 2 | Gait parameters of the enlarged perivascular spaces in the basal ganglia (BG-EPVS) group and the control group.

Gait parameters	BG-EPVS group	control group	*T*- or *Z*-value	*P*
Gait speed[a], m/min	41.4 ± 14.60	52.7 ± 12.84	6.059	< 0.001
Cadence[a], steps/min	99 ± 14.8	106 ± 13.1	2.170	< 0.001
Stride length[a], m	0.8 ± 0.21	1.0 ± 0.17	7.335	< 0.001
Stance phase percentage[a], %	63.6 ± 6.29	59.3 ± 7.68	-3.300	0.001
Foot fall[a], G	2.1 ± 0.80	3.0 ± 0.81	5.207	< 0.001
pre-swing angle[a], °	15.3 ± 6.49	22.6 ± 8.49	5.059	< 0.001
Tinetti gait test[b], sore	11 (9–12)	12 (12–12)	-7.255	< 0.001
Tinetti balance test[b], sore	15 (11–16)	16 (16–16)	-7.539	< 0.001

[a]Continuous variables with normal distribution were expressed as mean values ± standard deviation and were compared with Student t-test. The statistic value was T. [b]Continuous variables with non-normally distributions were expressed as median (interquartile range) and compared with the Wilcoxon rank-sum test. The statistic value was Z.

correlation between CSVD burden and motor performance in community-dwelling populations and found that EPVS was not associated with motor performance, a finding that differed from the results of our study. Su et al. (2017) classified patients with EPVS > 5 on the basal ganglia slice containing the maximum number into the EPVS group. The different results might be attributed to the different grouping methods and study populations.

Gait velocity is determined by both the stride length and gait cadence. In the present study, the gait velocity, stride length, and gait cadence were reduced in patients with large numbers of BG-EPVS. However, large numbers of BG-EPVS were only significantly related to gait velocity and stride length and not related to gait cadence after adjusting for WMH, lacunes, and CMB. We propose that stride length may be a more sensitive indicator than gait cadence. Aside from the lower velocity, we found that gait in subjects with severe BG-EPVS was characterized by an increased stance phase percentage, smaller pre-swing angle, reduced footfall, and lower Tinetti balance score, which indicated that severe BG-EPVS is related to poor balance.

TABLE 3 | Results of correlation analysis between gait performance and BG-EPVS.

Gait parameters	γ	*P*	γ (adjusted for confounders)	*P* (adjusted for confounders)
Gait speed	−0.367	< 0.001	−0.158	0.022
Cadence	−0.237	< 0.001	−0.133	0.055
Stride length	−0.431	< 0.001	−0.243	< 0.001
Stance phase percentage	0.269	< 0.001	0.170	0.014
Foot fall	−0.497	< 0.001	−0.313	< 0.001
pre-swing angle	−0.313	< 0.001	−0.166	0.016
Tinetti gait test	−0.449	< 0.001	−0.177	0.016
Tinetti balance test	−0.489	< 0.001	−0.238	0.001

The confounders included age, proportion of men, hypertension and stroke, level of blood creatinine, Fazekas score, the number of lacunae, and CMB.

TABLE 4 | Results of multiple linear regression analysis between gait performance and BG-EPVS.

Gait parameters	Model 1		Model 2		Model 3	
	β	*P*	β	*P*	β	*P*
Gait speed	−0.294	< 0.001	−0.281	< 0.001	−0.178	0.022
Cadence	−0.183	0.008	−0.180	0.015	−0.167	0.055
Stride length	−0.360	< 0.001	−0.343	< 0.001	−0.247	< 0.001
Stance phase percentage	0.233	0.001	0.241	0.001	0.205	0.014
Foot fall	−0.419	< 0.001	−0.404	< 0.001	−0.346	< 0.001
pre−swing angle	−0.203	0.002	−0.209	0.002	−0.191	0.016
Tinetti gait test	−0.381	< 0.001	−0.352	< 0.001	−0.188	0.016
Tinetti balance test	−0.419	< 0.001	−0.397	< 0.001	−0.255	0.001

Model 1: adjusted for age and sex. Model 2: model 1 + proportion of hypertension and stroke, level of blood creatinine. Model 3: model 2 + Fazekas score, the number of lacuna, and CMB.

The potential pathophysiological mechanisms underlying the association between BG-EPVS and gait are not completely understood. Perivascular spaces are thought to serve as a protolymphatic system and play an important role in maintaining neural homeostasis (Gouveia-Freitas and Bastos-Leite, 2021). Sulcal CSF is cleared through arachnoid granulations or it enters the parenchyma *via* the perivascular spaces, where it combines with interstitial fluid prior to exiting the brain. This perivascular drainage system also allows for the clearance of toxic metabolites within the parenchyma and possibly plays a role in the brain's immunological response (Iliff et al., 2012). EPVS are thought to be the result of perivascular blockages that disrupt the normal function of perivascular spaces. Although a few EPVS visible on MRI are normal, the presence of many is not normal, and they have been shown to be associated with some age-related disorders, including cognitive dysfunction, Parkinson's syndrome, WMH, and lacunar infarction (Rudie et al., 2018; Rundek et al., 2019; Gouveia-Freitas and Bastos-Leite, 2021). The basal ganglia–thalamocortical circuit, brain network efficiency, and loss of white matter microstructural integrity are involved in gait impairment in subjects with WMH, lacunes, and CMB (de Laat et al., 2011a; Cai et al., 2021). Recent findings showed that EPVS leads to white matter microstructural damage by causing a glymphatic dysfunction that results in the accumulation of toxic metabolic products. These metabolic products were harmful to the brain microenvironment (Che Mohd Nassir et al., 2021). Therefore, we speculated that large numbers of BG-EPVS may disrupt the functioning of the basal ganglia, brain network, and brain microenvironment and subsequently lead to gait disturbances and poor balance. The hypothesis should be tested and verified with multimodal MRI (e.g., functional MRI and diffusion tensor imaging) in the future.

Some limitations in the present study must be mentioned. First, our study was performed in a single center and the cohort may not represent the general population. Second, this was an observational study, and the causal relationship between BG-EPVS and gait disturbances cannot be established. Third, the mechanisms underlying the association between BG-EPVS and gait disturbances were not explored. A multicenter prospective cohort study using multimodal MRI to explore the mechanisms should be performed in the future. Despite these limitations, our study found that BG-EPVS was associated with gait disturbances in older individuals with CSVD with both quantitative and semiquantitative gait analysis. The novel findings provide some information about the consequences of BG-EPVS and risk factors for gait disturbances in older people.

In summary, we found that large numbers of BG-EPVS were independently related to gait disturbances and poor balance. Patients with severe BG-EPVS had lower gait speed, shorter stride length, longer stance phase, smaller pre-swing angle, and poorer foot control. The causal relationship and mechanisms should be further tested and explored in longitudinal studies in the future.

DATA AVAILABILITY STATEMENT

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Beijing Chaoyang Hospital Affiliated to Capital Medical University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

WH and SY conceived and designed the study. XL and SY participated in the screening participants and data collection. WQ and LY assessed the imagings. SY participated in the data analysis and drafted the manuscript. WH revised the manuscript. All authors read and approved the final manuscript to be published.

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 82001244).

REFERENCES

Abellan van Kan, G., Rolland, Y., Andrieu, S., Bauer, J., Beauchet, O., Bonnefoy, M., et al. (2009). Gait speed at usual pace as a predictor of adverse outcomes in community-dwelling older people an International Academy on Nutrition and Aging (IANA) Task Force. *J. Nutr. Health Aging* 13, 881–889. doi: 10.1007/s12603-009-0246-z

Baezner, H., Blahak, C., Poggesi, A., Pantoni, L., Inzitari, D., Chabriat, H., et al. (2008). Association of gait and balance disorders with age-related white matter changes: the LADIS study. *Neurology* 70, 935–942. doi: 10.1212/01.wnl.0000305959.46197.e6

Cai, M., Jacob, M. A., Norris, D. G., Duering, M., de Leeuw, F. E., and Tuladhar, A. M. (2021). Cognition mediates the relation between structural network efficiency and gait in small vessel disease. *Neuroimage Clin.* 30:102667. doi: 10.1016/j.nicl.2021.102667

Cannistraro, R. J., Badi, M., Eidelman, B. H., Dickson, D. W., Middlebrooks, E. H., and Meschia, J. F. (2019). CNS small vessel disease: a clinical review. *Neurology* 92, 1146–1156. doi: 10.1212/wnl.0000000000007654

Che Mohd Nassir, C. M. N., Damodaran, T., Yusof, S. R., Norazit, A., Chilla, G., et al. (2021). Aberrant Neurogliovascular Unit Dynamics in Cerebral Small Vessel Disease: A Rheological Clue to Vascular Parkinsonism. *Pharmaceutics* 13:1207. doi: 10.3390/pharmaceutics13081207

Chung, S. J., Yoo, H. S., Shin, N. Y., Park, Y. W., Lee, H. S., Hong, J. M., et al. (2021). Perivascular Spaces in the Basal Ganglia and Long-term Motor Prognosis in Newly Diagnosed Parkinson Disease. *Neurology* 96, e2121–e2131. doi: 10.1212/wnl.0000000000011797

de la Cámara, M., Higueras-Fresnillo, S., Martinez-Gomez, D., and Veiga, ÓL. (2019). Interday Reliability of the IDEEA Activity Monitor for Measuring Movement and Nonmovement Behaviors in Older Adults. *J. Aging Phys. Act* 27, 141–154. doi: 10.1123/japa.2017-0365

de Laat, K. F., Reid, A. T., Grim, D. C., Evans, A. C., Kötter, R., van Norden, A. G., et al. (2012). Cortical thickness is associated with gait disturbances in cerebral small vessel disease. *Neuroimage* 59, 1478–1484. doi: 10.1016/j.neuroimage.2011.08.005

de Laat, K. F., van den Berg, H. A., van Norden, A. G., Gons, R. A., Olde Rikkert, M. G., and de Leeuw, F. E. (2011b). Microbleeds are independently related to gait disturbances in elderly individuals with cerebral small vessel disease. *Stroke* 42, 494–497. doi: 10.1161/strokeaha.110.596122

de Laat, K. F., Tuladhar, A. M., van Norden, A. G., Norris, D. G., Zwiers, M. P., and de Leeuw, F. E. (2011a). Loss of white matter integrity is associated with gait disorders in cerebral small vessel disease. *Brain* 134, 73–83. doi: 10.1093/brain/awq343

de Laat, K. F., van Norden, A. G., Gons, R. A., van Oudheusden, L. J., van Uden, I. W., Bloem, B. R., et al. (2010). Gait in elderly with cerebral small vessel disease. *Stroke* 41, 1652–1658. doi: 10.1161/strokeaha.110.583229

Doubal, F. N., MacLullich, A. M., Ferguson, K. J., Dennis, M. S., and Wardlaw, J. M. (2010). Enlarged perivascular spaces on MRI are a feature of cerebral small vessel disease. *Stroke* 41, 450–454. doi: 10.1161/strokeaha.109.564914

Fazekas, F., Chawluk, J. B., Alavi, A., Hurtig, H. I., and Zimmerman, R. A. (1987). MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *AJR Am. J. Roentgenol.* 149, 351–356. doi: 10.2214/ajr.149.2.351

Gorelick, M. L., Bizzini, M., Maffiuletti, N. A., Munzinger, J. P., and Munzinger, U. (2009). Test-retest reliability of the IDEEA system in the quantification of step parameters during walking and stair climbing. *Clin. Physiol. Funct. Imag.* 29, 271–276. doi: 10.1111/j.1475-097X.2009.00864.x

Gouveia-Freitas, K., and Bastos-Leite, A. J. (2021). Perivascular spaces and brain waste clearance systems: relevance for neurodegenerative and cerebrovascular pathology. *Neuroradiology* 63, 1581–1597. doi: 10.1007/s00234-021-02718-7

Iliff, J. J., Wang, M., Liao, Y., Plogg, B. A., Peng, W., Gundersen, G. A., et al. (2012). A paravascular pathway facilitates CSF flow through the brain parenchyma and the clearance of interstitial solutes, including amyloid β. *Sci. Transl. Med.* 4:147ra111. doi: 10.1126/scitranslmed.3003748

Jie, W., Lin, G., Liu, Z., Zhou, H., Lin, L., Liang, G., et al. (2020). The Relationship Between Enlarged Perivascular Spaces and Cognitive Function: A Meta-Analysis of Observational Studies. *Front. Pharmacol.* 11:715. doi: 10.3389/fphar.2020.00715

Kim, Y. J., Kwon, H. K., Lee, J. M., Cho, H., Kim, H. J., Park, H. K., et al. (2016). Gray and white matter changes linking cerebral small vessel disease to gait disturbances. *Neurology* 86, 1199–1207. doi: 10.1212/wnl.0000000000002516

Litak, J., Mazurek, M., Kulesza, B., Szmygin, P., Litak, J., Kamieniak, P., et al. (2020). Cerebral Small Vessel Disease. *Int. J. Mol. Sci.* 21:9729. doi: 10.3390/ijms21249729

MacKinnon, C. D. (2018). Sensorimotor anatomy of gait, balance, and falls. *Handb. Clin. Neurol.* 159, 3–26. doi: 10.1016/b978-0-444-63916-5.00001-x

Rudie, J. D., Rauschecker, A. M., Nabavizadeh, S. A., and Mohan, S. (2018). Neuroimaging of Dilated Perivascular Spaces: From Benign and Pathologic Causes to Mimics. *J. Neuroimag.* 28, 139–149. doi: 10.1111/jon.12493

Rundek, T., Shpiner, D. S., and Margolesky, J. (2019). Perivascular spaces in basal ganglia - An innocent bystander in Parkinson's disease? *Mov. Disord* 34, 1585–1587. doi: 10.1002/mds.27824

Shin, N. Y., Park, Y. W., Yoo, S. W., Yoo, J. Y., Choi, Y., Jang, J., et al. (2021). Adverse effects of hypertension, supine hypertension, and perivascular space on cognition and motor function in PD. *NPJ Parkins. Dis.* 7:69. doi: 10.1038/s41531-021-00214-6

Stijntjes, M., de Craen, A. J., van der Grond, J., Meskers, C. G., Slagboom, P. E., and Maier, A. B. (2016). Cerebral Microbleeds and Lacunar Infarcts Are Associated with Walking Speed Independent of Cognitive Performance in Middle-Aged to Older Adults. *Gerontology* 62, 500–507. doi: 10.1159/000444583

Su, N., Zhai, F. F., Zhou, L. X., Ni, J., Yao, M., Li, M. L., et al. (2017). Cerebral Small Vessel Disease Burden Is Associated with Motor Performance of Lower and

Upper Extremities in Community-Dwelling Populations. *Front. Aging Neurosci.* 9:313. doi: 10.3389/fnagi.2017.00313

van der Holst, H. M., Tuladhar, A. M., Zerbi, V., van Uden, I. W. M., de Laat, K. F., van Leijsen, E. M. C., et al. (2018). White matter changes and gait decline in cerebral small vessel disease. *Neuroimag. Clin.* 17, 731–738. doi: 10.1016/j.nicl.2017.12.007

Verghese, J., LeValley, A., Hall, C. B., Katz, M. J., Ambrose, A. F., and Lipton, R. B. (2006). Epidemiology of gait disorders in community-residing older adults. *J. Am. Geriatr. Soc.* 54, 255–261. doi: 10.1111/j.1532-5415.2005.00580.x

Wagner, J. R., Schaper, M., Hamel, W., Westphal, M., Gerloff, C., Engel, A. K., et al. (2022). Combined Subthalamic and Nigral Stimulation Modulates Temporal Gait Coordination and Cortical Gait-Network Activity in Parkinson's Disease. *Front. Hum. Neurosci.* 16:812954. doi: 10.3389/fnhum.2022.812954

Wardlaw, J. M., Smith, E. E., Biessels, G. J., Cordonnier, C., Fazekas, F., Frayne, R., et al. (2013). Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol.* 12, 822–838. doi: 10.1016/s1474-4422(13)70124-8

Zhu, Y. C., Dufouil, C., Soumaré, A., Mazoyer, B., Chabriat, H., and Tzourio, C. (2010). High degree of dilated Virchow-Robin spaces on MRI is associated with increased risk of dementia. *J. Alz. Dis.* 22, 663–672. doi: 10.3233/jad-2010-100378

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

TYPE Systematic Review
PUBLISHED 19 July 2022
DOI 10.3389/fnagi.2022.878025

OPEN ACCESS

EDITED BY
Inga Liepelt-Scarfone,
University of Tübingen, Germany

REVIEWED BY
Walter Maetzler,
University of Kiel, Germany
Carla Tatiana Toro,
University of Warwick, United Kingdom
Ondrej Bezdicek,
Charles University, Czechia

*CORRESPONDENCE
Hao Zhang
crrczh2020@163.com

SPECIALTY SECTION
This article was submitted to Alzheimer's Disease and Related Dementias, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 17 February 2022
ACCEPTED 28 June 2022
PUBLISHED 19 July 2022

CITATION
Han K, Tang Z, Bai Z, Su W and Zhang H (2022) Effects of combined cognitive and physical intervention on enhancing cognition in older adults with and without mild cognitive impairment: A systematic review and meta-analysis.
Front. Aging Neurosci. 14:878025.
doi: 10.3389/fnagi.2022.878025

Effects of combined cognitive and physical intervention on enhancing cognition in older adults with and without mild cognitive impairment: A systematic review and meta-analysis

Kaiyue Han[1,2], Zhiqing Tang[1,2], Zirong Bai[3], Wenlong Su[1,2,4] and Hao Zhang[1,2,4,5]*

[1]School of Rehabilitation, Capital Medical University, Beijing, China, [2]China Rehabilitation Research Center, Beijing Bo'ai Hospital, Beijing, China, [3]Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Parkville, VIC, Australia, [4]University of Health and Rehabilitation Sciences, Qingdao, China, [5]Cheeloo College of Medicine, Shandong University, Jinan, China

Background: Combined cognitive and physical intervention is commonly used as a non-pharmacological therapy to improve cognitive function in older adults, but it is uncertain whether combined intervention can produce stronger cognitive gains than either single cognitive or sham intervention. To address this uncertainty, we performed a systematic review and meta-analysis to evaluate the effects of combined intervention on cognition in older adults with and without mild cognitive impairment (MCI).

Methods: We systematically searched eight databases for relevant articles published from inception to November 1, 2021. Randomized controlled trials (RCTs) and non-randomized controlled trials (NRCTs) were used to compare the effects of the combined intervention with a single cognitive or sham intervention on cognition in older adults with and without MCI aged $\geq$ 50 years. We also searched Google Scholar, references of the included articles, and relevant reviews. Two independent reviewers performed the article screening, data extraction, and bias assessment. GRADEpro was used to rate the strength of evidence, and RevMan software was used to perform the meta-analysis.

Results: Seventeen studies were included in the analysis, comprising eight studies of cognitively healthy older adults and nine studies of older adults with MCI. The meta-analysis showed that the combined intervention significantly improved most cognitive functions and depression (SMD = 0.99, 95% CI 0.54–1.43, $p < 0.0001$) in older adults compared to the control groups, but the intervention effects varied by cognition domains. However, there was no statistically significant difference in the maintenance between the combined and sham interventions (SMD = 1.34, 95% CI −0.58–3.27, $p = 0.17$). The subgroup analysis also showed that there was no statistical difference in the combined intervention to improve global cognition, memory, attention, and

executive function between cognitive healthy older adults and older adults with MCI.

Conclusions: Combined intervention improves cognitive functions in older adults with and without MCI, especially in global cognition, memory, and executive function. However, there was no statistical difference in the efficacy of the combined intervention to improve cognition between cognitive healthy older adults and older adults with MCI. Moreover, the maintenance of the combined intervention remains unclear due to the limited follow-up data and high heterogeneity. In the future, more stringent study designs with more follow-ups are needed further to explore the effects of combined intervention in older adults.

Systematic review registration: https://www.crd.york.ac.uk/PROSPERO/#recordDetails, identifier: CRD42021292490.

KEYWORDS

combined cognitive and physical intervention, cognition, older adults, mild cognitive impairment, systematic review, meta-analysis

Introduction

As the global population ages, cognitive decline has become an increasingly critical factor affecting the health and quality of life of older adults, ranging from normal cognitive function to mild cognitive impairment (MCI) even dementia (Anderson, 2020). In recent years, the prevalence of MCI has increased in older adults, exacerbating the potential impact on global physical and mental health (Vos et al., 2015; Overton et al., 2019). A study has shown that the proportion of participants with depression among older adults with MCI ranged from 20.1 to 44.3% (Panza et al., 2010), and improvement in this state of MCI plus depression (MCI/D) is an essential factor in improving quality of life. MCI is an early stage of memory loss or other cognitive ability loss in individuals who maintain the ability to independently perform most activities of daily living (ADL) (Jack et al., 2018). Moreover, MCI has a high risk of progressing into Alzheimer's disease (AD) and other dementias, with reported conversion rates of 50% in 2-3 years (Marioni et al., 2015) and even as high as 60–100% in 5–10 years (Albert et al., 2011).

MCI refers to a cognitive and functional decline syndrome with no currently available cure. At present, pharmacological treatments for patients with MCI have not been proven to be completely effective, and adverse effects have been observed (Briggs et al., 2016). Cognitive interventions using non-invasive and non-pharmacological treatments based on the theories of neuroplasticity (Greenwood and Parasuraman, 2010; Rajji, 2019) and rich environments have attracted more attention (Marlats et al., 2020; Liu et al., 2021). A previous study reported that older adults with and without MCI showed signs of cognitive decline to varying degrees, and combined cognitive and physical intervention effectively improves cognition (Wu et al., 2019), which also becomes a research hotspot in recent years. Shatil (2013) conducted a 16-week randomized controlled trial (RCT) of combined cognitive and physical intervention, single cognitive intervention, and sham intervention in 29, 33, and 29 cognitively healthy older adult subjects, respectively, and found that combined intervention was significantly better than single cognitive intervention in improving memory and naming, while sham intervention showed no improvement in cognition. Additionally, Park et al. (2019) conducted a 24-week RCT in 49 older adult subjects with amnesic MCI (aMCI), in which 25 subjects performed aerobic exercise while doing number crunching and found that combined intervention improved working memory and executive function, but the sham intervention did not improve cognition in the other 24 subjects.

Although many meta-analyses have reported the cognitive benefits of the combined intervention for older adults with and without MCI (Stanmore et al., 2017; Gheysen et al., 2018; Gavelin et al., 2021), they were mixed across age groups and included articles that varied considerably in terms of study designs, comparisons, and study qualities. Therefore, the efficacy of the combined intervention to improve cognition is yet to be determined, especially when compared to single cognitive intervention (Law et al., 2014; Wollesen et al., 2015; Zhu et al., 2016). To address the above limitations, this meta-analysis developed a more detailed inclusion criteria and separately reported the effects of the combined intervention compared with a single cognitive or sham intervention on cognition in older adults with and without MCI.

The objectives of this systematic review and meta-analysis are as follows: (1) to compare the effects of combined intervention with a single cognitive or sham intervention on cognition in older adults; (2) to explore the differences in

cognitive efficacy of the combined intervention for cognitively healthy older adults and those with MCI; and (3) to summarize and compare the maintenance and safety of combined intervention in order to provide practical strategies and methods for improving cognition in older adults.

Methods

We report the systematic review and meta-analysis following the Preferred Reporting Items for Systematic Reviews and Meta-analysis guidelines (Moher et al., 2009) and register the review in the International Prospective Register of Systematic Reviews (CRD42021292490).

Search strategy

We implemented the search strategy by using a combination of MESH terms, free-text words, and truncation retrieval, and we searched for articles on combined cognitive and physical intervention to enhance cognition in older adults with and without MCI published in PubMed, Embase, Web of Science, Cochrane Library, PsycINFO, Scopus, EBSCO and Ovid from inception to November 1, 2021. Furthermore, we screened all reference lists of the selected articles and related review articles, and we used the same search terms in Google Scholar to perform additional searches. The search was limited to publications in English. The complete search strategy (Supplementary Table S1) is provided in the Supplementary Material.

Selection criteria

The inclusion criteria of this meta-analysis is detailed below.

Participants

Studies were included if the participants: were cognitively healthy older adults or those diagnosed with MCI; had an age of 50 years or older.

Interventions

Combined cognitive and physical training as an intervention that is either a simultaneous or a sequential dual or multi-tasking (Gallou-Guyot et al., 2020), refers to performing two or even more cognitive and physical tasks separately or simultaneously (Tait et al., 2017; MacPherson, 2018). We did not limit the cognitive or physical training type in the combined intervention.

Comparisons

The intervention in the control group included either single cognitive or sham intervention (e.g., placebo control, blank control, and passive control) for older adults with or without MCI.

If the study had two or more control groups (e.g., single physical intervention, single cognitive intervention, or sham intervention), only data from the control group with single cognitive or sham intervention were included.

Outcomes

The primary outcome was cognitive function, including global cognitive function, memory, attention, and executive function; the secondary outcome was depression.

Cognition evaluation

Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA) to evaluate the global cognition; Logical Memory (LM), Digit Span Test (DST), Trail Making Test Parts A (TMT-A), Rey Auditory Verbal Learning Test (RAVLT), and Complex Figure Test (CFT) to evaluate memory function; Symbol Digit Substitution Test (SDST), Brief Test of Attention (BTA), Test of Everyday Attention (TEA), and attentional Matrices (AM) to evaluate attention; Trail Making Test Parts B (TMT-B) and Executive Function Cognitive Assessment Scale (FUCAS) to assess executive function; Stroop color-word test (SCWT) to evaluate inhibition and executive control function.

Depression evaluation

The included studies used the Geriatric Depression Scale (GDS) or the Cornell Scale for Depression in Dementia (CSDD) to assess depression.

Design

Studies that were randomized controlled trials (RCTs) or non-randomized controlled trials (NRCTs) were included in this review.

Study selection and data extraction

Two reviewers (HKY, TZQ) worked independently to screen the articles, extract information, and cross-check. In case of a disagreement, the articles were reviewed by a third reviewer (SWL). The authors of the original study were contacted *via* email to clarify or add any missing information. The articles were initially screened by reading the title and abstract before reading of the full text for re-screening. For each eligible study, we used a self-designed standardized form (Supplementary Table S2) to extract the first author's name, year of publication, country, clinical diagnosis of disease, number of participants, male ratio, age, education level, intervention methods, intervention characteristics, outcome measures, and drop-out.

Risk of bias and study quality assessment

Two reviewers (HKY, TZQ) independently assessed the studies according to the Cochrane Handbook for Systematic Reviews of Interventions (Higgins et al., 2011), and disagreements on assessments were resolved by discussion with the third reviewer (SWL). The assessment scale included the following seven items: random sequence generation and allocation concealment (selection bias), blinding of participants and personnel (performance bias), blinding of outcome assessment (detection bias), incomplete outcome data (attrition bias), selective reporting (reporting bias), and other bias. Three degrees of assessment were used to grade each item: "low," "unclear," and "high."

The PEDro scale, comprising 11 items, was used to assess the quality of the included studies, and studies with a score of seven or higher were considered to be of medium and high quality (Maher et al., 2003). Based on the risk of bias, inconsistency, indirectness, imprecision, and publication bias, the online GRADEpro method was used to evaluate the quality of evidence for pooled results in the meta-analysis (Cui et al., 2019).

Data analysis and statistical methods

We used RevMan software 5.4 to perform the meta-analysis. Since all data were continuous information and were pooled by the same outcome using inconsistent scales, we selected the Standardized Mean Difference (SMD) as an effective indicator and provided the 95% confidence interval (CI). We used the Cochrane Q statistic to qualitatively determine whether heterogeneity existed among the included studies (test level $\alpha = 0.05$), while the I^2 statistic was used to quantitatively determine the magnitude of heterogeneity. If the P-value was ≥ 0.1 and $I^2 \leq 50\%$, the heterogeneity was considered to be insignificant and we selected the fixed-effects (FE) model. Conversely, if the heterogeneity was considered to be significant, we selected the random-effects (RE) model and performed a subgroup analysis and sensitivity analysis to identify the factors that contributed to the heterogeneity. Descriptive analysis was performed if the source of heterogeneity could not ultimately be determined.

Results

Study selection

The flowchart of study selection is shown in Figure 1. We initially retrieved 1,353 articles from the nine databases and identified one article through other sources. Eight hundred and forty-four articles remained after removing duplicates using a reference management software. After reading the titles and abstracts for screening, 797 articles were excluded. Subsequently, after screening the full text of the remaining 47 articles, 10 articles were not full data available, nine articles had a non-conforming control group, three articles had no cognitive assessment results, six articles had no conforming neuropsychological tests, and the full text of two articles were not available. Finally, 17 articles were included in this review.

Characteristics of the included studies

As shown in Supplementary Table S2, eight studies of cognitively healthy older adults were eligible (Fabre et al., 2002; Marmeleira et al., 2009; Shatil, 2013; Hars et al., 2014; Nishiguchi et al., 2015; Rahe et al., 2015a,b; Morita et al., 2018), with 181 participants in the combined intervention group, 68 in the single cognitive intervention group, and 142 in the sham intervention group. Regarding the study design, six studies were RCTs (Fabre et al., 2002; Marmeleira et al., 2009; Shatil, 2013; Hars et al., 2014; Nishiguchi et al., 2015; Rahe et al., 2015a) and two studies were NRCTs (Rahe et al., 2015b; Morita et al., 2018). Regarding the modes of combined intervention, four studies performed simultaneous combined cognitive and physical training (Marmeleira et al., 2009; Shatil, 2013; Hars et al., 2014; Nishiguchi et al., 2015; Morita et al., 2018) and four studies performed sequential combined intervention (Fabre et al., 2002; Shatil, 2013; Rahe et al., 2015a,b), all of which reported greater cognitive gains in the combined intervention. Regarding the comparison condition, three studies used single cognitive intervention (Shatil, 2013; Rahe et al., 2015a,b), one study used reading as a placebo control (Shatil, 2013), four studies used a blank control (Fabre et al., 2002; Marmeleira et al., 2009; Hars et al., 2014; Nishiguchi et al., 2015), and one study used non-exercise as a passive control (Morita et al., 2018). Additionally, five studies implemented interventions longer than 12 weeks (Marmeleira et al., 2009; Shatil, 2013; Hars et al., 2014; Nishiguchi et al., 2015; Morita et al., 2018). Only one study had a follow-up - up to 1 year -and reported that combined intervention can produce more significant long-term effects than single cognitive intervention, especially in attention (Rahe et al., 2015b).

Nine studies of older adults with MCI were eligible (Kounti et al., 2011; Lam et al., 2015; Delbroek et al., 2017; Park, 2017; Donnezan et al., 2018; Mrakic-Sposta et al., 2018; Park et al., 2019, 2020; Rojasavastera et al., 2020), with 217 participants in the combined intervention group, 41 in the single cognitive intervention group, and 176 in the sham intervention group. Regarding the study design, eight studies were RCTs (Lam et al., 2015; Delbroek et al., 2017; Park, 2017; Donnezan et al., 2018; Mrakic-Sposta et al., 2018; Park et al., 2019, 2020; Rojasavastera et al., 2020) and one study was NRCT (Kounti et al., 2011). Regarding the modes of combined intervention, seven studies included simultaneous combined cognitive and physical

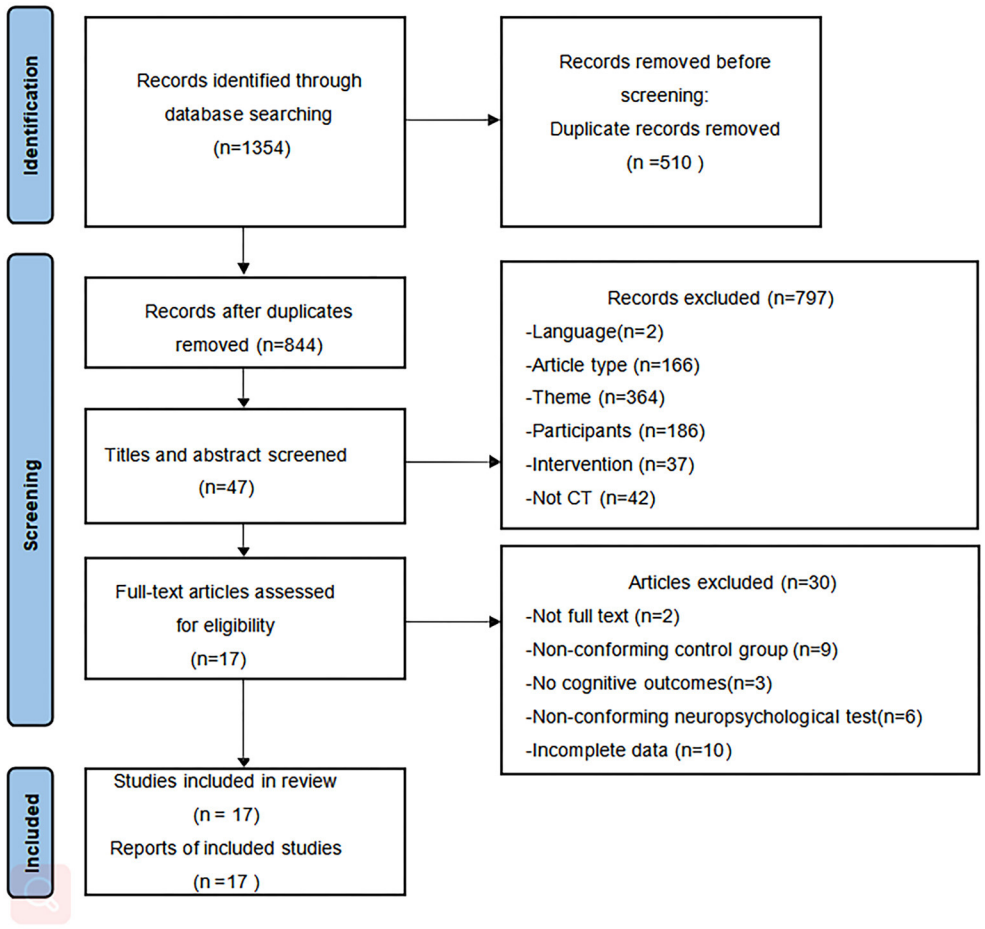


FIGURE 1
PRISMA flowchart of study selection.

training (Kounti et al., 2011; Delbroek et al., 2017; Park, 2017; Donnezan et al., 2018; Mrakic-Sposta et al., 2018; Park et al., 2019, 2020) and two studies performed sequential combined intervention (Lam et al., 2015; Rojasavastera et al., 2020), all of which reported greater cognitive improvements in the combined intervention. Regarding the comparison condition, three studies used single cognitive intervention (Park, 2017; Donnezan et al., 2018; Park et al., 2020), one study used social activities as a placebo control (Lam et al., 2015), and five studies used a blank control (Kounti et al., 2011; Delbroek et al., 2017; Mrakic-Sposta et al., 2018; Park et al., 2019; Rojasavastera et al., 2020). Additionally, four studies implemented interventions longer than 12 weeks (Kounti et al., 2011; Lam et al., 2015; Donnezan et al., 2018; Park et al., 2019). Only three studies had follow-up—up to 1, 3, and 6 months, respectively—and they also reported greater long-term cognitive improvements in combined intervention group (Donnezan et al., 2018; Park et al., 2019; Rojasavastera et al., 2020).

Risk of bias and quality assessment

The PEDro scale showed that all studies were non-low quality (Supplementary Table S2). The risk of bias of the included studies is shown in Figure 2. Of the 17 studies included, three studies did not use randomization methods (Kounti et al., 2011; Rahe et al., 2015b; Morita et al., 2018) and four did not report allocation concealment (Marmeleira et al., 2009; Kounti et al., 2011; Morita et al., 2018; Park et al., 2019). The participants and personnel of three studies were not blinded to the combined intervention because of the intervention design's characteristics, which were considered to have a high risk of bias (Park, 2017; Donnezan et al., 2018; Park et al., 2019), while the outcome assessments of seven studies were blinded (Kounti et al., 2011; Hars et al., 2014; Lam et al., 2015; Rahe et al., 2015a; Delbroek et al., 2017; Park, 2017; Morita et al., 2018). A total of 13 studies showed a low risk of bias in attrition, reporting, and other biases (Fabre et al., 2002; Marmeleira et al., 2009; Shatil, 2013; Hars

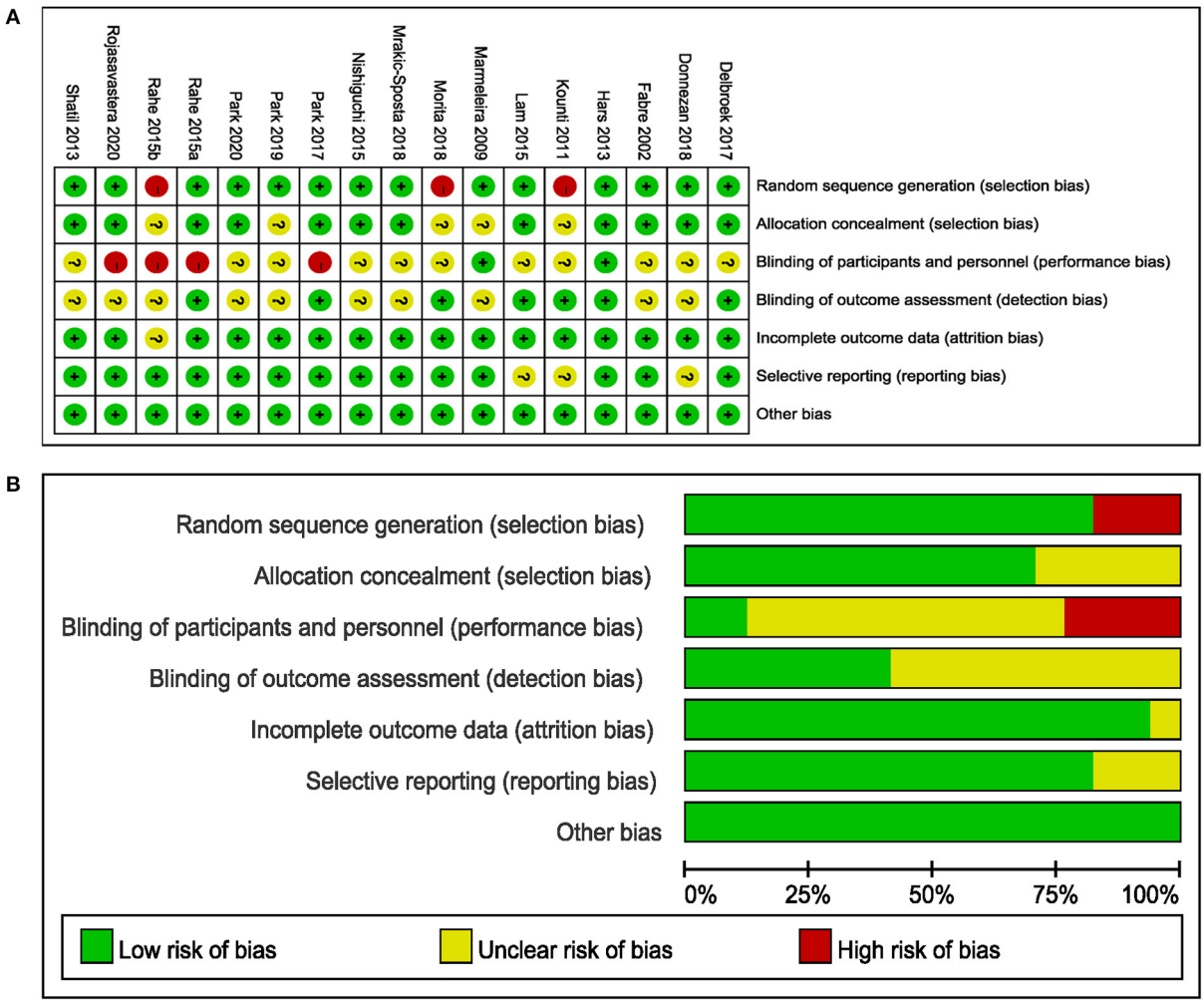


FIGURE 2
Results from the Cochrane risk of bias (ROB) tool. **(A)** ROB graph and **(B)** ROB summary.

et al., 2014; Nishiguchi et al., 2015; Rahe et al., 2015a; Delbroek et al., 2017; Park, 2017; Morita et al., 2018; Mrakic-Sposta et al., 2018; Park et al., 2019, 2020; Rojasavastera et al., 2020).

For global cognitive function, the GRADE ratings from the included studies showed the effectiveness of "moderate" and "low" using the MMSE and MoCA to measure outcome (Table 1).

Effects of the combined intervention

Effects of combined intervention in cognitively healthy older adults

Global cognition

Three studies used MMSE to assess the efficacy of the combined intervention on global cognition in cognitively healthy older adults (Hars et al., 2014; Nishiguchi et al., 2015; Morita et al., 2018). Application of the RE model to the pooled SMD revealed that the global cognitive level was significantly higher in the combined group than in the control group (SMD = 1.77, 95% CI 0.94–2.59, $p < 0.0001$, Supplementary Figure S1). Next, due to the high heterogeneity ($I^2 = 73\%$, $\chi^2 = 7.53$, $p = 0.02$), we excluded one study at a time to perform sensitivity analysis. The result after excluding one study (Morita et al., 2018) showed the heterogeneity decreased ($I^2 = 53\%$, $\chi^2 = 2.11$, $p = 0.15$), as well as a change in the overall pooled effect (SMD = 1.40, 95% CI 0.85–1.96, $p < 0.00001$, Figure 3).

Cognition domains

Based on different cognition domains, we performed a subgroup analysis that compared the efficacy of the combined intervention with single cognitive, sham interventions to improve cognition in cognitively healthy older adults. Compared with single cognitive intervention (Figure 4A), the

TABLE 1 Summary of the GRADEpro.

Question: Effects of combined intervention in the global cognition for older adults with MCI.

Setting: Hospitals in mainland China

Intervention: combined group

Comparison: control group

Outcome measure	No of studies	No of the participants	Anticipated absolute effects* (95% CI)	Certainty of the evidence (GRADE)
MMSE	4	305	SMD 0.81 higher (0.51 higher to 1.11 higher)	⊕⊕⊕◯ Moderate[a]
MoCA	4	95	SMD 0.93 higher (0.12 lower to 1.98 higher)	⊕⊕◯◯ Low[a,b]

Certainty of the evidence (GRADE)

High: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate: We are moderately confident in the effect estimate; The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect.

Very low: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect.

*The risk in the intervention group (and its 95% CI) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).
CI, confidence interval; MMSE, the Mini-Mental State Examination; SMD, standardized mean difference; MoCA, the Montreal Cognitive Assessment.
[a]Most of the RCTs were low quality with an inadequate level of blinding and unclear risk of concealment of allocation.
[b]The statistical test for heterogeneity showed that large variation ($I^2 > 50\%$) existed in point estimates due to the among study differences.

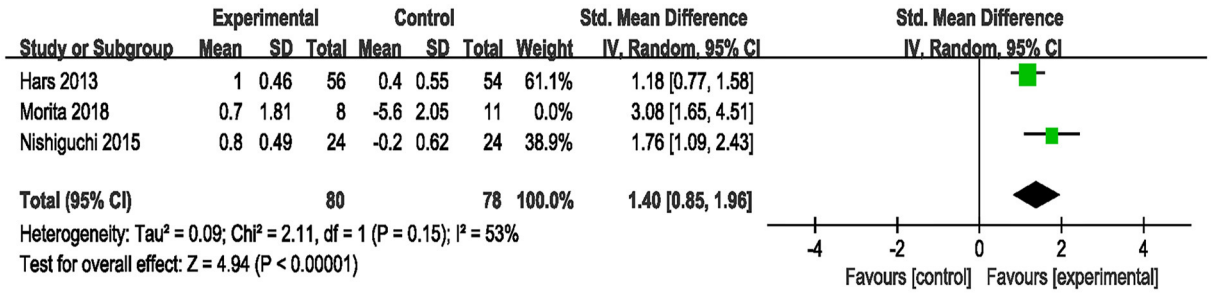


FIGURE 3
Forest plot of the efficacy of the combined intervention on global cognition in cognitively healthy older adults compared to the control group (sensitive analysis).

pooled SMD showed that combined intervention significantly improved working memory (SMD = 0.45, 95% CI 0.06–0.84, $p = 0.02$), but no significant improvement in figural memory (SMD = 0.57, 95% CI −0.14–1.28, $p = 0.11$) and inhibition (SMD = 0.78, 95% CI −0.01–1.57, $p = 0.05$). Compared with the sham intervention (Figure 4B), the combined intervention significantly improved memory recall (SMD = 1.93, 95% CI 1.33–2.54, $p < 0.00001$), divided attention (SMD = 1.01, 95% CI 0.14–1.87, $p = 0.02$) and speed processing (SMD = 1.91, 95% CI 0.79–3.03, $p = 0.0008$). However, this subgroup analysis showed a significant heterogeneity ($I^2 = 75\%$, $\chi^2 = 20.40$, $p = 0.001$), and we did not perform sensitivity analysis to identify the heterogeneity sources because of the limited number of studies in each subgroup. Different cognitive rating scales, intervention frequency, and duration may have contributed to the observed heterogeneity.

Effects of combined intervention in older adults with MCI

Global cognition

Eight studies assessed the efficacy of the combined intervention on global cognition using the MMSE and MoCA (Kounti et al., 2011; Lam et al., 2015; Delbroek et al., 2017; Park, 2017; Mrakic-Sposta et al., 2018; Park et al., 2019, 2020; Rojasavastera et al., 2020). In a subgroup analysis based

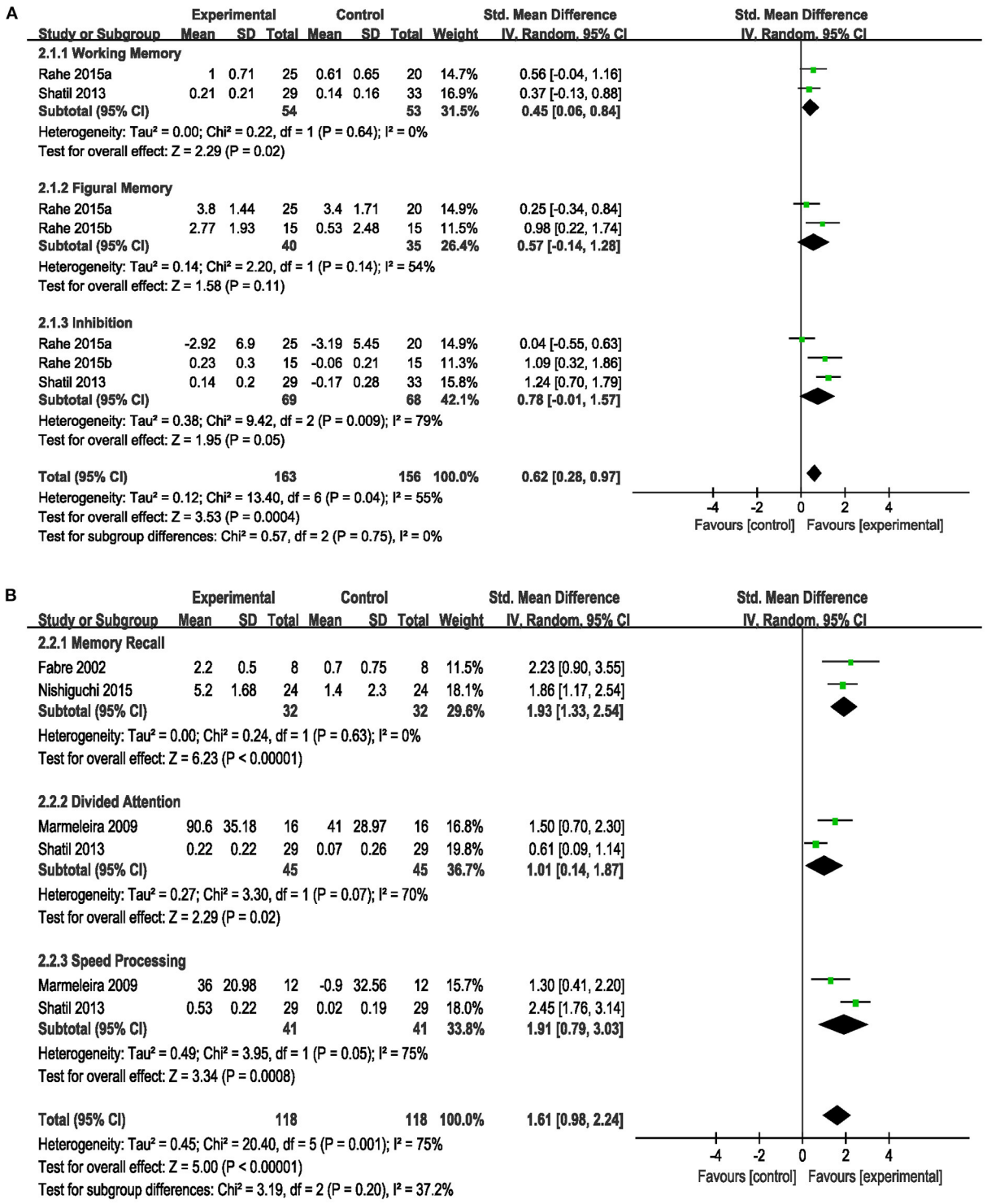


FIGURE 4
Forest plot of the efficacy of the combined intervention on cognition domains in cognitively healthy older adults. **(A)** Combined intervention vs. single cognitive intervention, **(B)** combined intervention vs. sham intervention.

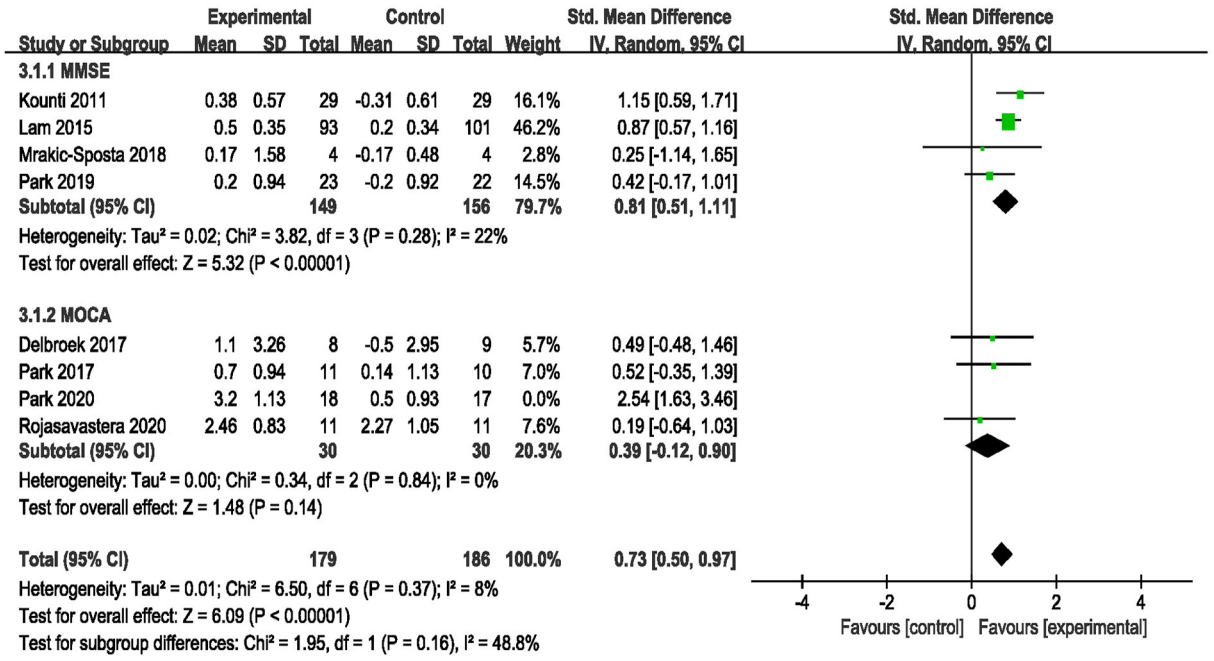


FIGURE 5
Forest plot of the efficacy of the combined intervention on global cognition in older adults with MCI compared to the control group (sensitive analysis).

on different cognitive scales, the pooled SMD showed that combined intervention was more beneficial for improving global cognition (SMD = 0.83, 95% CI 0.41–1.25, $p = 0.0001$, Supplementary Figure S2). We performed a sensitivity analysis due to the high heterogeneity ($I^2 = 66\%$, $\chi^2 = 20.39$, $p = 0.005$). After excluding one study (Park et al., 2020), the heterogeneity decreased ($I^2 = 8\%$, $\chi^2 = 6.50$, $p = 0.37$), and the pooled result also changed (SMD = 0.73, 95% CI 0.50–0.97, $P < 0.00001$, Figure 5).

Cognition domains

Subgroup analysis compared the efficacy of the combined intervention with single cognitive, sham intervention to improve cognition in older adults with MCI. Compared with the single cognitive intervention (Supplementary Figure S3), the results showed that combined intervention significantly improved working memory (SMD = 2.00, 95% CI 0.40–3.60, $p = 0.01$) and speed processing (SMD = 3.98, 95% CI 2.78–5.17, $p < 0.00001$). When we performed a sensitivity analysis due to the high heterogeneity ($I^2 = 90\%$, $\chi^2 = 29.43$, $p < 0.00001$), the heterogeneity decreased ($I^2 = 57\%$, $\chi^2 = 2.34$, $p = 0.13$) after excluding one study (Park et al., 2020), and the overall pooled effect in working memory also changed (SMD = 1.18, 95% CI 0.29–2.07, $p = 0.009$, Figure 6A). Additionally, compared with the sham intervention (Figure 6B), under acceptable heterogeneity ($I^2 = 54\%$, $\chi^2 = 10.90$, $p = 0.05$), the subgroup analysis revealed that combined intervention significantly improved memory recall (SMD = 0.97, 95% CI 0.67–1.26, $p < 0.00001$) and executive function (SMD = 1.77, 95% CI 1.31–2.23, $p < 0.00001$), but no significant improvement in attention (SMD = 0.96, 95% CI −0.10–2.02, $p = 0.08$).

Depression

Only three studies assessed the efficacy of the combined intervention to improve depression in older adults with MCI, with one study using CSDD (Lam et al., 2015) and two studies using GDS (Park, 2017; Park et al., 2019). Under acceptable heterogeneity ($I^2 = 48\%$, $\chi^2 = 3.84$, $p = 0.15$), the pooled results showed that combined intervention induced a significant improvement in depression (SMD = 0.99, 95% CI 0.54–1.43, $p < 0.0001$, Figure 7).

Efficacy differences of combined intervention between cognitively healthy older adults and older adults with MCI

As shown in Table 2, in order to reduce heterogeneity, we used the same comparison and outcome assessment scales to analyze the efficacy differences of the combined intervention in older adults with and without MCI. Therefore, the number of studies included was limited. After sensitivity analysis, the subgroup analysis showed that there were no statistical difference within the combined intervention to improve global cognition (SMD = 1.40, 95% CI 0.85–1.96, $p < 0.00001$; vs.

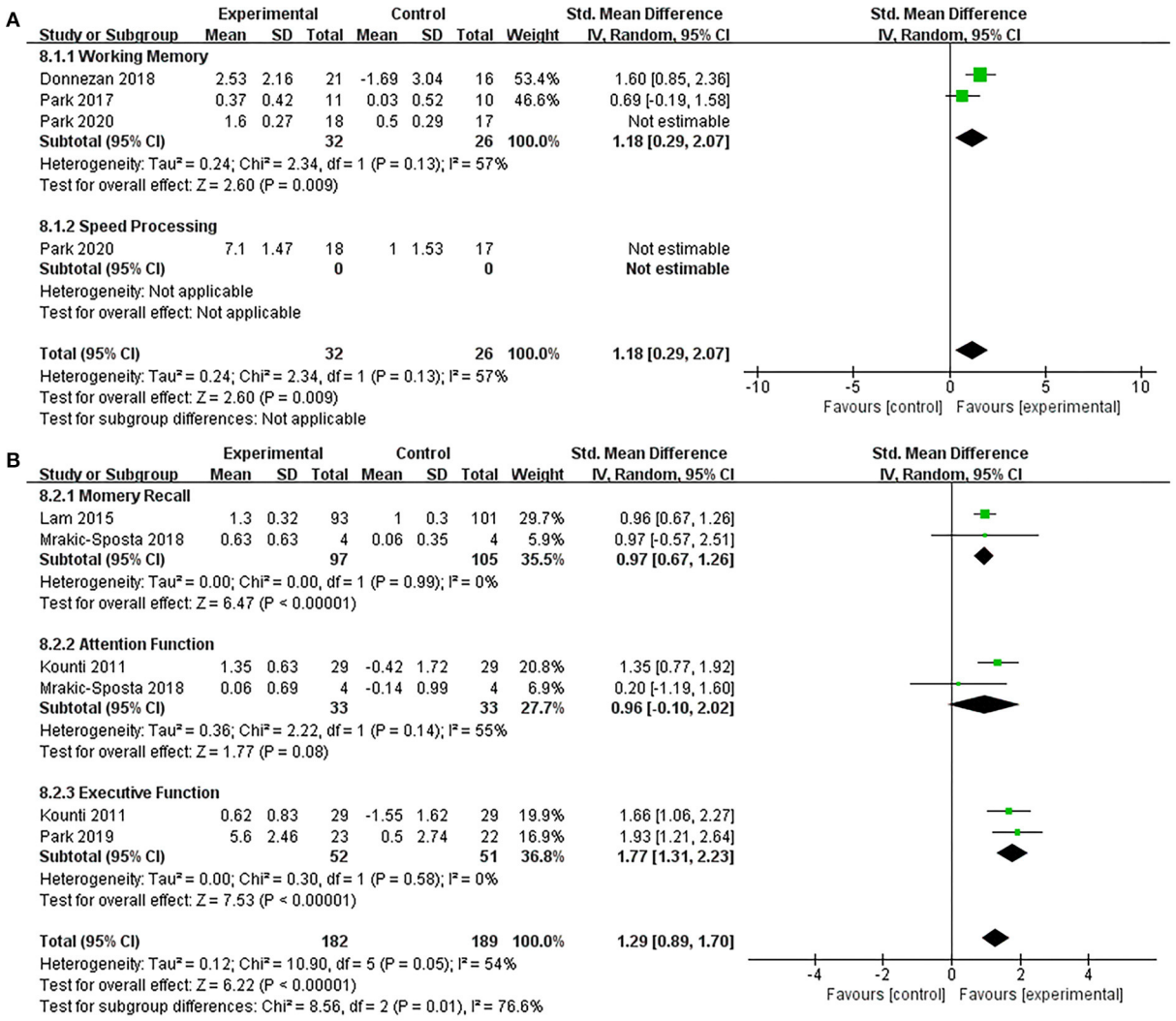


FIGURE 6
Forest plot of the efficacy of the combined intervention on cognition domains in older adults with MCI. **(A)** Combined intervention vs. single cognitive intervention (sensitive analysis), **(B)** combined intervention vs. sham intervention.

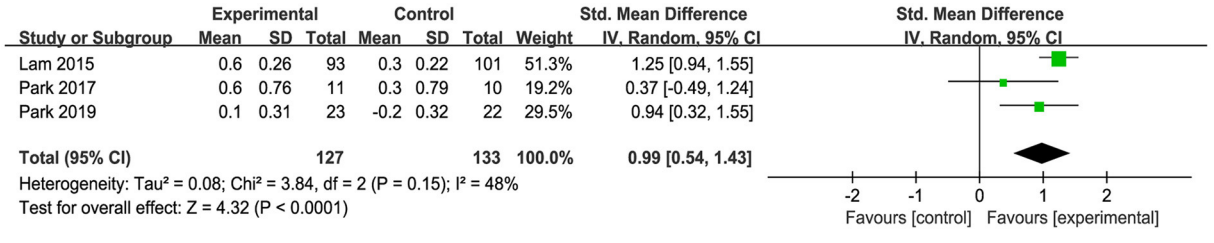


FIGURE 7
Forest plot of the efficacy of the combined intervention on depression in older adults with MCI compared with the control group.

TABLE 2 Efficacy differences of combined intervention on cognition between cognitively healthy older adults and older adults with MCI after sensitive analysis.

Outcomes	Subgroup	No. of studies	SMD 95% CI	Homogeneity				Test for overall effect		Test for subgroup[1,2] differences	
				Q	*df*	*p*	I^2,%	*Z*	*p*	*p*	I^2,%
Global cognition[a]	Subgroup[1]	3	1.40	0.85 to 1.96	2.11	1	0.15	53	4.94	<0.00001	0.07	70.2
	Subgroup[2]	4	0.81	0.51 to 1.11	3.82	3	0.28	22	5.32	<0.00001		
Memory[b]	Subgroup[1]	2	0.70	0.18 to 1.23	0.89	1	0.34	0	2.61	0.009	0.36	0
	Subgroup[2]	3	1.18	0.29 to 2.07	2.34	1	0.13	57	2.60	0.009		
Attention[c]	Subgroup[1]	1	−0.04	−0.60 to 0.51	NA	NA	NA	NA	0.15	0.88	0.94	0
	Subgroup[2]	2	−0.08	−0.94 to 0.78	NA	NA	NA	NA	0.19	0.85		
Executive function[d]	Subgroup[1]	2	0.39	−0.42 to 1.20	NA	NA	NA	NA	0.94	0.35	0.67	0
	Subgroup[2]	2	0.62	−0.07 to 1.30	1.20	1	0.27	16	1.77	0.08		

SMD, standardized mean difference; CI, confidence interval; Subgroup[1], the cognitively healthy older adults group; Subgroup[2], the older adults with MCI group; NA, not applicable.
[a]Results of a study excluded after sensitivity analysis (Morita et al., 2018).
[b]Results of a study excluded after sensitivity analysis (Park et al., 2020).
[c]Results of a study excluded after sensitivity analysis (Donnezan et al., 2018).
[d]Results of a study excluded after sensitivity analysis (Nishiguchi et al., 2015).

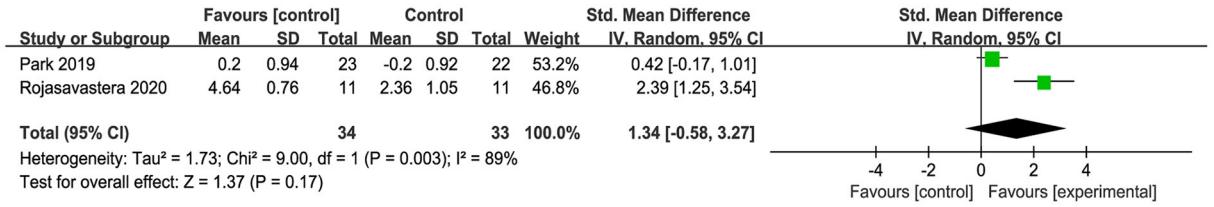


FIGURE 8
Forest plot of the maintenance on global cognition in older adults with MCI compared with sham intervention.

SMD = 0.81, 95% CI 0.51–1.11, $p < 0.00001$), memory (SMD = 0.70, 95% CI 0.18–1.23, $p = 0.009$; vs. SMD = 1.18, 95% CI 0.29–2.07, $p = 0.009$), attention (SMD = −0.04, 95% CI −0.60–0.51, $p = 0.88$; vs. SMD = −0.08, 95% CI −0.94–0.78, $p = 0.85$), and executive function (SMD = 0.39, 95% CI −0.42–1.20, $p = 0.35$; vs. SMD = 0.62, 95% CI −0.07–1.30, $p = 0.08$) between cognitive healthy older adults and older adults with MCI.

The maintenance and safety of combined intervention

As shown in Figure 8, only two studies were included to assess the maintenance of the combined intervention on global cognition in older adults with MCI compared to the sham intervention (Park et al., 2019; Rojasavastera et al., 2020), and the results showed no statistical difference (SMD = 1.34, 95% CI −0.58–3.27, $p = 0.17$). Similarly, due to limited follow-up data, we did not perform a subgroup analysis based on the different cognitive scales, which may have been a source of the observed high heterogeneity.

The minor adverse event was the risk of falls in older adults while performing physical training. The researchers increased safety protection and education for older adults to minimize this risk.

Moderator analysis for combined intervention

As shown in Table 3, because the outcome assessment scales and comparisons were not fully the same among studies, we only assessed the effect of the moderator variables on the efficacy of the combined intervention in order to improve global cognition in older adults with MCI. The results of the subgroup analyses showed that age (SMD = 0.73, 95% CI −0.21–1.66, $p = 0.13$; vs. SMD = 0.74, 95% CI 0.49–0.99, $p < 0.00001$), education (SMD = 0.75, 95% CI 0.49–1.01, $p < 0.00001$; vs. SMD = 0.73, 95% CI −0.21–1.66, $p = 0.13$), intervention duration (SMD = 0.37, 95% CI −0.1–0.85, $p = 0.13$; vs. SMD = 0.79, 95% CI 0.08–1.511, $p = 0.03$) and the mode of combined intervention (SMD = 0.69, 95% CI 0.35–1.03, $p < 0.0001$; vs. SMD = 0.65, 95% CI 0.03–1.27, $p = 0.04$) had an effect on the efficacy of the combined intervention in

TABLE 3 Effects of moderators on the efficacy of combined intervention to improve cognition in older adults with MCI after sensitive analysis.

Moderator variable	Level (subgroup)	No. of studies	SMD	95% CI	Homogeneity				Test for overall effect		Test for subgroup differences	
					Q	*df*	*p*	I^2, %	*Z*	*p*	*p*	I^2, %
Age[a,c]	≤70 years	2	0.73	−0.21 to 1.66	3.49	1	0.06	71	1.52	0.13	0.97	0
	>70years	5	0.74	0.49 to 0.99	2.54	3	0.47	0	5.81	<0.00001		
Education[b,c]	Elementary school	4	0.75	0.49 to 1.01	2.04	2	0.36	2	5.67	<0.00001	0.96	0
	Middle to high school	2	0.73	−0.21 to 1.66	3.49	1	0.06	71	1.52	0.13		
Intervention duration[c]	≤3 months	5	0.37	−0.11 to 0.85	0.37	3	0.95	0	1.52	0.13	0.23	32.5
	3–6 months	2	0.79	0.08 to 1.51	3.10	1	0.08	68	2.18	0.03		
	>6 months	1	0.87	0.57 to 1.16	NA	NA	NA	NA	5.76	<0.00001		
Mode of combined intervention[c]	Simultaneous	6	0.69	0.35 to 1.03	4.11	4	0.39	3	3.99	<0.0001	0.90	0
	Sequential	2	0.65	0.03 to 1.27	2.21	1	0.14	55	2.06	0.04		

SMD, standardized mean difference; CI, confidence interval; NA, not applicable.
[a]One study was excluded because the mean age of participants was not reported (Park, 2017).
[b]Two studies was excluded because education level was not reported (Delbroek et al., 2017; Mrakic-Sposta et al., 2018).
[c]Results of a study excluded after sensitivity analysis (Park et al., 2020).

improving cognition. However, we were unable to draw a precise conclusion about whether intervention frequency affected the efficacy of the combined intervention because there was only one study with an intervention frequency more than 3 days per week.

Discussion

Summary of findings

Global cognition

The results of our analysis showed that the combined intervention group was superior to the control group in improving global cognition in older adults with and without MCI, which is consistent with the results of other studies (Karssemeijer et al., 2017; Gavelin et al., 2021). Dual or multi-tasking training of combined cognitive and physical intervention is the basis to improve global cognition and ADL, which can reduce neurophysiological changes in cognition by reducing bilateral prefrontal cortical oxygenation, increasing hippocampal volume, and increasing white matter integrity (Tait et al., 2017). However, due to the limited number of studies, we did not perform subgroup analyzes according to different comparison conditions in global cognition. Additionally, seven studies assessed global cognition by MMSE (Kounti et al., 2011; Hars et al., 2014; Lam et al., 2015; Nishiguchi et al., 2015; Delbroek et al., 2017; Park, 2017; Morita et al., 2018), but two of them (Lam et al., 2015; Morita et al., 2018) using modified MMSE, which may limit the credibility of the results, so the results should be interpreted carefully. This also emphasizes the necessity on further evaluate the specific cognition domains to draw accurate conclusions.

Cognition domains

There is growing evidence that even the aging brain displays cognitive plasticity (Park and Bischof, 2013; Pauwels et al., 2018). Yang et al. (2020) reported that combined intervention improved most cognitive function in older adults with and without MCI, but had no effect on attention, and it was uncertain whether these positive effects would persist (Yang et al., 2020), which is consistent with our findings. Based on the theory of dual-task interference, the superior effect of the combined intervention may not be observed in the short term because of the cognitive and physical interaction. Therefore, the follow-up assessments are critical when studying the efficacy of the combined intervention to improve cognition in older adults in the future.

Depression

Based on the pathophysiological mechanisms of cognitive deficits and depression, we found an apparent correlation between them (Geda et al., 2006; Pellegrino et al., 2013), In older adults with MCI, patients with depression ranged from 20.1 to 44.3% (Panza et al., 2010). The statistical results of a study showed a positive correlation between the severity of depression and MCI, with depression significantly affecting delayed recall, verbal fluency, attention, and executive function in older adults (Dillon et al., 2009). Furthermore, depression as a risk factor for MCI has significant public health implications. Our results revealed that combined intervention had a small to moderate positive effect on depression, and other studies have reported that improvements in depression reduce the severity of MCI (Kessing et al., 2011; Pellegrino et al., 2013). A study by Barnes and Yaffe (2011) reported that a 10% reduction in depression prevalence could lead to 326,000 fewer AD cases worldwide.

Efficacy differences of combined intervention between cognitively healthy older adults and older adults with MCI

Our review reported that there was no statistical difference in the efficacy of the combined intervention for improving cognition in older adults with and without MCI, which is inconsistent with the findings of Wu et al. (2019), who suggested that the combined intervention was more effective in improving global cognition in older adults with MCI compared to cognitively healthy older adults (Wu et al., 2019). We used the same comparison and outcome assessment scales to assess efficacy differences, resulting in a limited number of studies included for this outcome; therefore, the results should be interpreted cautiously.

The maintenance and safety of combined intervention

Due to limited follow-up data, this meta-analysis only reported that the efficacy of the combined intervention in improving global cognition in older adults with MCI was not maintained (Park et al., 2019; Rojasavastera et al., 2020); however, another three studies found positive maintenance of the combined intervention (Barnes et al., 2013; Lee et al., 2016; Norouzi et al., 2019). In summary, we found heterogeneity primarily in two areas: the types of physical tasks within the combined intervention and the modes of the combined intervention. Regarding the types of physical task, resistance training (Norouzi et al., 2019), combined aerobic and resistance training (Barnes et al., 2013; Lee et al., 2016) improved the long-term working memory and global cognition within older adults with MCI; however, aerobic training alone was not found to have positive efficacy maintenance (Park et al., 2019; Rojasavastera et al., 2020). Thus far, combined aerobic and resistance training is the most commonly used and effective type of exercise (Kelly et al., 2014). Furthermore, the modes of combined intervention are divided into sequential (Park et al., 2019; Rojasavastera et al., 2020) and simultaneous interventions (Barnes et al., 2013; Norouzi et al., 2019). It was found that simultaneous intervention is superior to sequential intervention during efficacy maintenance, which may be based upon the mechanisms of physical-cognitive interaction. This result validates the intervention mode derived in our review as an influential factor in the efficacy of the combined intervention and is also consistent with the results of other meta-analyses (Zhu et al., 2016). However, it remains controversial whether the time of each sequential intervention is the same as that of simultaneous intervention (Joubert and Chainay, 2018).

Except for a slight risk of falls, none of the included studies reported significant adverse events during the combined intervention. Furthermore, due to the limited sample size, the safety and maintenance of the combined intervention will need to be validated *via* multicenter studies with larger sample sizes, and more follow-ups.

Moderators analysis for combined intervention

In terms of demographic characteristics, this review found that age and education level were influential factors in the efficacy of the combined intervention. Moreover, the combined intervention was more effective during advanced age as well as less educated older adults, which may be related to this population's lower baseline cognitive performance. Previous studies found a positive association between age and the efficacy of the combined intervention, while no correlation was reported in education (Powers et al., 2013; Toril et al., 2014; Qarni and Salardini, 2019).

Different intervention durations also affected the efficacy of the combined intervention. Law et al. (2014) found that an intervention duration of 3–6 months was more beneficial for improving cognition in older adults with MCI (Law et al., 2014), and is consistent with the results of our study. Suzuki et al. (2012) also reported that a 6-month combined intervention effectively improved cognition in older adults; however, the efficacy did not last until the end of the 12 month treatment regimen. Due to the limited number of included studies, we were unable to draw a precise conclusion about whether intervention frequency affected the efficacy of the combined intervention. However, a previous meta-analysis found that high-frequency combined intervention might be ineffective (Zhu et al., 2016). Two studies on working memory also reported that high-frequency intervention might lead to cognitive fatigue causing participants to drop out of the study (Penner et al., 2012; Wang et al., 2014). In conclusion, selecting the appropriate intervention frequency and duration is likely to be an essential factor in improving the efficacy of a combined intervention.

Limitations

This meta-analysis also has some limitations. First, the number of included studies was limited. Second, the outcome measurements did not use imaging, electroencephalogram (EEG), or other objective evaluation methods. The evidence suggests structural and functional magnetic resonance imaging or electrophysiological measurements of brain activity can more accurately evaluate the changes of specific areas in the brain (Bherer et al., 2013). Third, only English articles were included.

Implications for future studies

Two points need to be improved in the future. First, to maximize the effect of intervention, future studies need

to stringently design the mode, frequency, and duration of the combined intervention, and a long-term follow-up. Second, we need to select more appropriate outcome measurement indexes, comprehensive neuropsychological assessments, and objective evaluation tools (e.g., imaging and EEG) to accurately assess the efficacy of the combined intervention.

Conclusion

In summary, this meta-analysis showed that combined cognitive and physical intervention effectively improves cognition in older adults with and without MCI compared with single cognitive or sham intervention, although the intervention effects vary by cognition domains. However, it is challenging to draw an obvious conclusion in the combined intervention maintenance because of the limitations. Additionally, there was no statistical difference in the efficacy of the combined intervention to improve cognition between cognitive healthy older adults and older adults with MCI. The results should be interpreted carefully due to the different intervention designs and the diversity of evaluation methods. In the future, more stringent study designs with more follow-ups are needed to clarify the effects of the combined intervention and provide guidance on the optimum intervention regime for improving cognitive function in older adults.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author/s.

Author contributions

KH contributed to study design, literature search, figures, data extraction, data analysis, and writing. ZT contributed to literature search, data extraction, and data analysis. ZB and WS contributed to figures, data extraction, data interpretation, and writing. HZ contributed to study design and data interpretation. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Key Research and Development Program of China (Grant No. 2018YFC2001703), Capital Health Research and Development of Special Fund (Grant No. 2020-1-6011), and China Rehabilitation Research Center Key Project (Grant No. 2021ZX-02).

Acknowledgments

We are grateful to those who offered any help in this article. The authors thank AiMi Academic Services (www.aimieditor.com) for the English language editing and review services.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.878025/full#supplementary-material

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., et al. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement. J. Alzheimers Assoc.* 7, 270–279. doi: 10.1016/j.jalz.2011.03.008

Anderson, N. D. (2020). State of the science on mild cognitive impairment. *J. Gerontol. B. Psychol. Sci. Soc. Sci.* 75, 1359–1360. doi: 10.1093/geronb/gbaa040

Barnes, D. E., Santos-Modesitt, W., Poelke, G., Kramer, A. F., Castro, C., Middleton, L. E., et al. (2013). The mental activity and exercise (MAX) trial: a randomized controlled trial to enhance cognitive function in older adults. *JAMA Intern. Med.* 173, 797–804. doi: 10.1001/jamainternmed.2013.189

Barnes, D. E., and Yaffe, K. (2011). The projected effect of risk factor reduction on Alzheimer's disease prevalence. *Lancet Neurol.* 10, 819–828. doi: 10.1016/S1474-4422(11)70072-2

Bherer, L., Erickson, K. I., and Liu-Ambrose, T. (2013). A review of the effects of physical activity and exercise on cognitive and brain functions in older adults. *J. Aging Res.* 2013, 657508. doi: 10.1155/2013/657508

Briggs, R., Kennelly, S. P., and O'Neill, D. (2016). Drug treatments in Alzheimer's disease. *Clin. Med. Lond. Engl.* 16, 247–253. doi: 10.7861/clinmedicine.16-3-247

Cui, S., Chen, N., Yang, M., Guo, J., Zhou, M., Zhu, C., et al. (2019). Cerebrolysin for vascular dementia. *Cochrane Database Syst. Rev.* 2019, CD008900. doi: 10.1002/14651858.CD008900.pub3

Delbroek, T., Vermeylen, W., and Spildooren, J. (2017). The effect of cognitive-motor dual task training with the biorescue force platform on cognition, balance and dual task performance in institutionalized older adults: a randomized controlled trial. *J. Phys. Ther. Sci.* 29, 1137–1143. doi: 10.1589/jpts.29.1137

Dillon, C., Allegri, R. F., Serrano, C. M., Iturry, M., Salgado, P., Glaser, F. B., et al. (2009). Late- versus early-onset geriatric depression in a memory research center. *Neuropsychiatr. Dis. Treat.* 5, 517–526. doi: 10.2147/NDT.S7320

Donnezan, L., Perrot, A., Belleville, S., Bloch, F., and Kemoun, G. (2018). Effects of simultaneous aerobic and cognitive training on executive functions, cardiovascular fitness and functional abilities in older adults with mild cognitive impairment. *Ment. Health Phys. Act.* 15, 78–87. doi: 10.1016/j.mhpa.2018.06.001

Fabre, C., Chamari, K., Mucci, P., Massé-Biron, J., and Préfaut, C. (2002). Improvement of cognitive function by mental and/or individualized aerobic training in healthy elderly subjects. *Int. J. Sports Med.* 23, 415–421. doi: 10.1055/s-2002-33735

Gallou-Guyot, M., Mandigout, S., Bherer, L., and Perrochon, A. (2020). Effects of exergames and cognitive-motor dual-task training on cognitive, physical and dual-task functions in cognitively healthy older adults: an overview. *Ageing Res. Rev.* 63, 101135. doi: 10.1016/j.arr.2020.101135

Gavelin, H. M., Dong, C., Minkov, R., Bahar-Fuchs, A., Ellis, K. A., Lautenschlager, N. T., et al. (2021). Combined physical and cognitive training for older adults with and without cognitive impairment: a systematic review and network meta-analysis of randomized controlled trials. *Ageing Res. Rev.* 66, 101232. doi: 10.1016/j.arr.2020.101232

Geda, Y. E., Knopman, D. S., Mrazek, D. A., Jicha, G. A., Smith, G. E., Negash, S., et al. (2006). Depression, apolipoprotein E genotype, and the incidence of mild cognitive impairment: a prospective cohort study. *Arch. Neurol.* 63, 435–440. doi: 10.1001/archneur.63.3.435

Gheysen, F., Poppe, L., DeSmet, A., Swinnen, S., Cardon, G., De Bourdeaudhuij, I., et al. (2018). Physical activity to improve cognition in older adults: can physical activity programs enriched with cognitive challenges enhance the effects? A systematic review and meta-analysis. *Int. J. Behav. Nutr. Phys. Act.* 15, 63. doi: 10.1186/s12966-018-0697-x

Greenwood, P. M., and Parasuraman, R. (2010). Neuronal and cognitive plasticity: a neurocognitive framework for ameliorating cognitive aging. *Front. Aging Neurosci.* 2, 150. doi: 10.3389/fnagi.2010.00150

Hars, M., Herrmann, F. R., Gold, G., Rizzoli, R., and Trombetti, A. (2014). Effect of music-based multitask training on cognition and mood in older adults. *Age Ageing* 43, 196–200. doi: 10.1093/ageing/aft163

Higgins, J. P. T., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., et al. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 343, d5928. doi: 10.1136/bmj.d5928

Jack, C. R., Bennett, D. A., Blennow, K., Carrillo, M. C., Dunn, B., Haeberlein, S. B., et al. (2018). NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement. J. Alzheimers Assoc.* 14, 535–562. doi: 10.1016/j.jalz.2018.02.018

Joubert, C., and Chainay, H. (2018). Aging brain: the effect of combined cognitive and physical training on cognition as compared to cognitive and physical training alone - a systematic review. *Clin. Interv. Aging* 13, 1267–1301. doi: 10.2147/CIA.S165399

Karssemeijer, E. G. A., Aaronson, J. A., Bossers, W. J., Smits, T., Olde Rikkert, M. G. M., and 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 Res. Rev.* 40, 75–83. doi: 10.1016/j.arr.2017.09.003

Kelly, M. E., Loughrey, D., Lawlor, B. A., Robertson, I. H., Walsh, C., and Brennan, S. (2014). The impact of exercise on the cognitive functioning of healthy older adults: a systematic review and meta-analysis. *Ageing Res. Rev.* 16, 12–31. doi: 10.1016/j.arr.2014.05.002

Kessing, L. V., Forman, J. L., and Andersen, P. K. (2011). Do continued antidepressants protect against dementia in patients with severe depressive disorder? *Int. Clin. Psychopharmacol.* 26, 316–322. doi: 10.1097/YIC.0b013e32834ace0f

Kounti, F., Bakoglidou, E., Agogiatou, C., Lombardo, N. B. E., Serper, L. L., and Tsolaki, M. (2011). RHEA,* a nonpharmacological cognitive training intervention in patients with mild cognitive impairment: a pilot study. *Top. Geriatr. Rehabil.* 27, 289–300. doi: 10.1097/TGR.0b013e31821e59a9

Lam, L. C.-W., Chan, W. C., Leung, T., Fung, A. W.-T., and Leung, E. M.-F. (2015). Would older adults with mild cognitive impairment adhere to and benefit from a structured lifestyle activity intervention to enhance cognition?: a cluster randomized controlled trial. *PLoS ONE* 10, e0118173. doi: 10.1371/journal.pone.0118173

Law, L. L. F., Barnett, F., Yau, M. K., and Gray, M. A. (2014). Effects of combined cognitive and exercise interventions on cognition in older adults with and without cognitive impairment: a systematic review. *Ageing Res. Rev.* 15, 61–75. doi: 10.1016/j.arr.2014.02.008

Lee, Y.-Y., Wu, C.-Y., Teng, C.-H., Hsu, W.-C., Chang, K.-C., and Chen, P. (2016). Evolving methods to combine cognitive and physical training for individuals with mild cognitive impairment: study protocol for a randomized controlled study. *Trials* 17, 526. doi: 10.1186/s13063-016-1650-4

Liu, C. S., Herrmann, N., Song, B. X., Ba, J., Gallagher, D., Oh, P. I., et al. (2021). Exercise priming with transcranial direct current stimulation: a study protocol for a randomized, parallel-design, sham-controlled trial in mild cognitive impairment and Alzheimer's disease. *BMC Geriatr.* 21, 677. doi: 10.1186/s12877-021-02636-6

MacPherson, S. E. (2018). Definition: dual-tasking and multitasking. *Cortex J. Devoted Study Nerv. Syst. Behav.* 106, 313–314. doi: 10.1016/j.cortex.2018.06.009

Maher, C. G., Sherrington, C., Herbert, R. D., Moseley, A. M., and Elkins, M. (2003). Reliability of the PEDro scale for rating quality of randomized controlled trials. *Phys. Ther.* 83, 713–721. doi: 10.1093/ptj/83.8.713

Marioni, R. E., Proust-Lima, C., Amieva, H., Brayne, C., Matthews, F. E., Dartigues, J.-F., et al. (2015). Social activity, cognitive decline and dementia risk: a 20-year prospective cohort study. *BMC Public Health* 15, 1089. doi: 10.1186/s12889-015-2426-6

Marlats, F., Bao, G., Chevallier, S., Boubaya, M., Djabelkhir-Jemmi, L., Wu, Y.-H., et al. (2020). SMR/theta neurofeedback training improves cognitive performance and eeg activity in elderly with mild cognitive impairment: a pilot study. *Front. Aging Neurosci.* 12, 147. doi: 10.3389/fnagi.2020.00147

Marmeleira, J. F., Godinho, M. B., and Fernandes, O. M. (2009). The effects of an exercise program on several abilities associated with driving performance in older adults. *Accid. Anal. Prev.* 41, 90–97. doi: 10.1016/j.aap.2008.09.008

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 339, b2535. doi: 10.1136/bmj.b2535

Morita, E., Yokoyama, H., Imai, D., Takeda, R., Ota, A., Kawai, E., et al. (2018). Effects of 2-year cognitive-motor dual-task training on cognitive function and motor ability in healthy elderly people: a pilot study. *Brain Sci.* 8, E86. doi: 10.3390/brainsci8050086

Mrakic-Sposta, S., Di Santo, S. G., Franchini, F., Arlati, S., Zangiacomi, A., Greci, L., et al. (2018). Effects of combined physical and cognitive virtual reality-based training on cognitive impairment and oxidative stress in mci patients: a pilot study. *Front. Aging Neurosci.* 10, 282. doi: 10.3389/fnagi.2018.00282

Nishiguchi, S., Yamada, M., Tanigawa, T., Sekiyama, K., Kawagoe, T., Suzuki, M., et al. (2015). A 12-week physical and cognitive exercise program can improve cognitive function and neural efficiency in community-dwelling older adults: a randomized controlled trial. *J. Am. Geriatr. Soc.* 63, 1355–1363. doi: 10.1111/jgs.13481

Norouzi, E., Vaezmosavi, M., Gerber, M., Pühse, U., and Brand, S. (2019). Dual-task training on cognition and resistance training improved both balance and working memory in older people. *Phys. Sportsmed.* 47, 471–478. doi: 10.1080/00913847.2019.1623996

Overton, M., Pihlsgård, M., and Elmståhl, S. (2019). Prevalence and incidence of mild cognitive impairment across subtypes, age, and sex. *Dement. Geriatr. Cogn. Disord.* 47, 219–232. doi: 10.1159/000499763

Panza, F., Frisardi, V., Capurso, C., D'Introno, A., Colacicco, A. M., Imbimbo, B. P., et al. (2010). Late-life depression, mild cognitive impairment, and dementia: possible continuum? *Am. J. Geriatr. Psychiatry* 18, 98–116. doi: 10.1097/JGP.0b013e3181b0fa13

Park, D. C., and Bischof, G. N. (2013). The aging mind: neuroplasticity in response to cognitive training. *Dialogues Clin. Neurosci.* 15, 109–119. doi: 10.31887/DCNS.2013.15.1/dpark

Park, H., Park, J. H., Na, H. R., Hiroyuki, S., Kim, G. M., Jung, M. K., et al. (2019). Combined intervention of physical activity, aerobic exercise, and cognitive exercise intervention to prevent cognitive decline for patients with mild cognitive impairment: a randomized controlled clinical study. *J. Clin. Med.* 8, E940. doi: 10.3390/jcm8070940

Park, J. H. (2017). Does a cognitive-exercise combined dual-task training have better clinical outcomes for the elderly people with mild cognitive impairment than a single-task training? *Ther. Sci. Rehabil.* 6, 71–83. doi: 10.22683/tsnr.2017.6.2.071

Park, J. S., Jung, Y.-J., and Lee, G. (2020). Virtual reality-based cognitive-motor rehabilitation in older adults with mild cognitive impairment: a randomized controlled study on motivation and cognitive function. *Healthc. Basel Switz.* 8, E335. doi: 10.3390/healthcare8030335

Pauwels, L., Chalavi, S., and Swinnen, S. P. (2018). Aging and brain plasticity. *Aging* 10, 1789–1790. doi: 10.18632/aging.101514

Pellegrino, L. D., Peters, M. E., Lyketsos, C. G., and Marano, C. M. (2013). Depression in cognitive impairment. *Curr. Psychiatry Rep.* 15, 384. doi: 10.1007/s11920-013-0384-1

Penner, I.-K., Vogt, A., Stöcklin, M., Gschwind, L., Opwis, K., and Calabrese, P. (2012). Computerised working memory training in healthy adults: a comparison of two different training schedules. *Neuropsychol. Rehabil.* 22, 716–733. doi: 10.1080/09602011.2012.686883

Powers, K. L., Brooks, P. J., Aldrich, N. J., Palladino, M. A., and Alfieri, L. (2013). Effects of video-game play on information processing: a meta-analytic investigation. *Psychon. Bull. Rev.* 20, 1055–1079. doi: 10.3758/s13423-013-0418-z

Qarni, T., and Salardini, A. (2019). A multifactor approach to mild cognitive impairment. *Semin. Neurol.* 39, 179–187. doi: 10.1055/s-0039-1678585

Rahe, J., Becker, J., Fink, G. R., Kessler, J., Kukolja, J., Rahn, A., et al. (2015a). Cognitive training with and without additional physical activity in healthy older adults: cognitive effects, neurobiological mechanisms, and prediction of training success. *Front. Aging Neurosci.* 7, 187. doi: 10.3389/fnagi.2015.00187

Rahe, J., Petrelli, A., Kaesberg, S., Fink, G. R., Kessler, J., and Kalbe, E. (2015b). Effects of cognitive training with additional physical activity compared to pure cognitive training in healthy older adults. *Clin. Interv. Aging* 10, 297–310. doi: 10.2147/CIA.S74071

Rajji, T. K. (2019). Impaired brain plasticity as a potential therapeutic target for treatment and prevention of dementia. *Expert Opin. Ther. Targets* 23, 21–28. doi: 10.1080/14728222.2019.1550074

Rojasavastera, R., Bovonsunthonchai, S., Hiengkaew, V., and Senanarong, V. (2020). Action observation combined with gait training to improve gait and cognition in elderly with mild cognitive impairment A randomized controlled trial. *Dement. Neuropsychol.* 14, 118–127. doi: 10.1590/1980-57642020dn14-020004

Shatil, E. (2013). Does combined cognitive training and physical activity training enhance cognitive abilities more than either alone? A four-condition randomized controlled trial among healthy older adults. *Front. Aging Neurosci.* 5, 8. doi: 10.3389/fnagi.2013.00008

Stanmore, E., Stubbs, B., Vancampfort, D., de Bruin, E. D., and Firth, J. (2017). The effect of active video games on cognitive functioning in clinical and non-clinical populations: a meta-analysis of randomized controlled trials. *Neurosci. Biobehav. Rev.* 78, 34–43. doi: 10.1016/j.neubiorev.2017.04.011

Suzuki, T., Shimada, H., Makizako, H., Doi, T., Yoshida, D., Tsutsumimoto, K., et al. (2012). Effects of multicomponent exercise on cognitive function in older adults with amnestic mild cognitive impairment: a randomized controlled trial. *BMC Neurol.* 12, 128. doi: 10.1186/1471-2377-12-128

Tait, J. L., Duckham, R. L., Milte, C. M., Main, L. C., and Daly, R. M. (2017). Influence of sequential vs. simultaneous dual-task exercise training on cognitive function in older adults. *Front. Aging Neurosci.* 9, 368. doi: 10.3389/fnagi.2017.00368

Toril, P., Reales, J. M., and Ballesteros, S. (2014). Video game training enhances cognition of older adults: a meta-analytic study. *Psychol. Aging* 29, 706–716. doi: 10.1037/a0037507

Vos, S. J. B., Verhey, F., Frölich, L., Kornhuber, J., Wiltfang, J., Maier, W., et al. (2015). Prevalence and prognosis of Alzheimer's disease at the mild cognitive impairment stage. *Brain J. Neurol.* 138, 1327–1338. doi: 10.1093/brain/awv029

Wang, Z., Zhou, R., and Shah, P. (2014). Spaced cognitive training promotes training transfer. *Front. Hum. Neurosci.* 8, 217. doi: 10.3389/fnhum.2014.00217

Wollesen, B., Voelcker-Rehage, C., Willer, J., Zech, A., and Mattes, K. (2015). Feasibility study of dual-task-managing training to improve gait performance of older adults. *Aging Clin. Exp. Res.* 27, 447–455. doi: 10.1007/s40520-014-0301-4

Wu, C., Yi, Q., Zheng, X., Cui, S., Chen, B., Lu, L., et al. (2019). Effects of mind-body exercises on cognitive function in older adults: a meta-analysis. *J. Am. Geriatr. Soc.* 67, 749–758. doi: 10.1111/jgs.15714

Yang, C., Moore, A., Mpofu, E., Dorstyn, D., Li, Q., and Yin, C. (2020). Effectiveness of combined cognitive and physical interventions to enhance functioning in older adults with mild cognitive impairment: a systematic review of randomized controlled trials. *Gerontologist* 60, 633–642. doi: 10.1093/geront/gnz149

Zhu, X., Yin, S., Lang, M., He, R., and Li, J. (2016). The more the better? A meta-analysis on effects of combined cognitive and physical intervention on cognition in healthy older adults. *Ageing Res. Rev.* 31, 67–79. doi: 10.1016/j.arr.2016.07.003

frontiers | Frontiers in Aging Neuroscience

ORIGINAL RESEARCH
published: 22 July 2022
doi: 10.3389/fnagi.2022.829544

Instrumental Activities of Daily Living by Subjective and Objective Measures: The Impact of Depression and Personality

Katya Numbers[1], Sujin Jang[1], Henry Brodaty[1], Perminder S. Sachdev[1], Brian Draper[1] and Simone Reppermund[1,2]*

[1] Centre for Healthy Brain Ageing, Discipline of Psychiatry and Mental Health, Faculty of Medicine and Health, University of New South Wales, Sydney, NSW, Australia, [2] Department of Developmental Disability Neuropsychiatry, Faculty of Medicine and Health, School of Psychiatry, University of New South Wales, Sydney, NSW, Australia

Objective: Previous research shows that depression and personality are independently associated with self- and informant-reports of the ability to perform instrumental activities of daily living (IADLs). However, less is known about the association between depression and personality and performance-based measures of IADLs. We aimed to determine how depression and personality predict self-and informant-reports of IADL compared to performance-based measures of IADLs in a sample of older adults with normal cognition (NC) and Mild Cognitive Impairment (MCI).

Methods: Participants consisted of 385 older adults with NC (n = 235), or a diagnosis of MCI (n = 150), aged between 76 and 99-years from the Sydney Memory and Ageing Study. Participants underwent comprehensive neuropsychological and clinical assessments to determine global cognition and clinical diagnoses. Personality traits were measured by the NEO Five-Factor Inventory (NEO-FFI) and depression by the Geriatric Depression Scale (GDS). Subjective IADLs were self- and informant-reported Bayer Activities of Daily Living (B-ADL) scales and objective IADL was the Sydney Test of Activities of Daily Living in Memory Disorders (STAM). Linear regressions examined the relationship between depression and personality and the three types of IADL measures, controlling for all covariates and global cognition.

Results: Participant-reported IADL, although associated with global cognition, was more strongly associated with GDS and NEO-FFI scores (conscientiousness and neuroticism). Informant-reported IADL was strongly associated with both global cognition and participants' GDS scores. STAM scores were not associated with participants' GDS or NEO-FFI scores; instead, they were predicted by demographics and global cognition.

Conclusion: These results suggest that performance-based measures of IADL may provide more objective and reliable insight into an individual's underlying functional ability and are less impacted by the participants' mood and personality compared

OPEN ACCESS

Edited by:
Inga Liepelt-Scarfone, University of Tübingen, Germany

Reviewed by:
Edmond Teng, Genentech, Inc., United States
Cândida Alves, Federal University of Maranhão, Brazil

****Correspondence:***
Katya Numbers
k.numbers@unsw.edu.au

Specialty section:
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

Received: *05 December 2021*
Accepted: *21 June 2022*
Published: *22 July 2022*

Citation:
Numbers K, Jang S, Brodaty H, Sachdev PS, Draper B and Reppermund S (2022) Instrumental Activities of Daily Living by Subjective and Objective Measures: The Impact of Depression and Personality. Front. Aging Neurosci. 14:829544. doi: 10.3389/fnagi.2022.829544

to subjectively reported IADL. We argue that performance-based IADL measures are preferable when trying to accurately assess everyday functional ability and its relationship to cognitive status. Where performance-based measures are not available (e.g., in some clinical settings), informant ratings should be sought as they are less influenced by the participant's personality and mood compared to self-reports.

Keywords: IADL, functional ability, depression, openness, neuroticism, conscientiousness

INTRODUCTION

The cornerstone of functional independence among older adults is an intact ability to perform necessary activities of daily living (ADL). The loss of independence in these activities is a key factor affecting the quality of life in individuals with dementia and their caregivers (Kempen et al., 1997). Impairment in ADL is a key feature in the diagnosis of dementia, with loss of ability to perform basic activities of daily living (BADL), or well-rehearsed everyday tasks such as ambulating, toileting, bathing, grooming, and feeding (Mlinac and Feng, 2016), distinguishing mild dementia from moderate or severe dementia according to the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (Mitchell and Miller, 2008). Instrumental activities of daily living (IADL), such as driving, shopping, and managing finances or medications (Gold, 2012), recruit multiple cognitive domains and require planning and cognitive flexibility to complete (Mitchell and Miller, 2008). IADL is mostly preserved in Mild Cognitive Impairment (MCI) (Teng et al., 2010) but becomes impaired to varying degrees in mild dementia. The reliable and accurate assessment of IADL in an individual with cognitive deficits is therefore critical for determining a dementia diagnosis.

Currently, the most common method for assessing IADL is self- or informant- (i.e., close friend/family member) reported questionnaires (Tabert et al., 2002), which are quick and easy to collect and require relatively few resources or training to administer. Evidence suggests that, when given the opportunity for multiple observations, subjectively reported IADL gives a reasonably accurate representation of real-world performance (Schmitter-Edgecombe et al., 2011). However, a downside to self-reported measures of IADL is that they are based on people's perceptions of their own functioning, which may result in overestimation or underestimation of actual ability (Coman and Richardson, 2006), especially as poorer awareness of IADL difficulties is associated with cognitive impairment (Albert et al., 1999; Steward et al., 2019). Further, self-reported IADL is known to be affected by an individual's mood and personality traits. Specifically, individuals experiencing depressive symptoms (Taş et al., 2007; Karakurt and Unsal, 2013; de Paula et al., 2015; Storeng et al., 2018) tend to overreport impairments in IADL, as do those who score higher on neuroticism (Krueger et al., 2006) and lower on conscientiousness (Suchy et al., 2010) according to a five-factor personality model [e.g., NEO-Five Factor Inventory (Costa, 1992)]. Informant-reported IADL is not subject to many of these limitations and therefore is more frequently used to assess participants' functional capacity in research and clinical settings. However, informant-reported IADL can also be impacted by the participant's depressive symptoms or personality traits, as some studies have shown (Votruba et al., 2015). Other factors such as informants' own depressive symptoms and personality (Argüelles et al., 2001; Pfeifer et al., 2013), perceived burden of caring for the participant (Zanetti et al., 1999), social desirability and halo effects (Pereira et al., 2010), and limited insight into the daily routines of the participant (Martyr et al., 2014), can influence informant-reports of IADL as well.

To provide a more objective alternative to subjective self- and informant-reported IADL, performance-based measurements have been developed (Zanetti et al., 1999). Performance-based measures require participants to perform various IADL activities, such as measuring out medications or counting money for shopping, under direct observation from the assessor (Sikkes and Rotrou, 2014). Performance-based measures are less subject to bias, lack of insight, and the informant's knowledge of and feelings toward the individual (Zanetti et al., 1999; Griffith et al., 2003; Goldberg et al., 2010). Therefore, performance-based IADLs are more sensitive to subtle decrements in IADL function (Goldberg et al., 2010; Pereira et al., 2010). However, performance-based measures are more time-consuming and expensive (Moore et al., 2007) and require specialized materials and training to administer (Reppermund et al., 2017). Moreover, it is possible that an individual's level of depression and personality factors may influence their performance on objective measures of IADL as well.

While some studies have examined the association between personality and self-report vs. performance-based IADL (Suchy et al., 2010), to our knowledge, no study has investigated the relationship between depressive symptoms and personality and scores on a self-report, informant-report, and performance-based measure of IADL concurrently. It is important to understand this relationship clearly to determine whether interventions that target depression, a modifiable condition, may help to preserve functional capacity in older adults without a dementia diagnosis (Albert et al., 1999). Moreover, it is necessary to clarify how this relationship differs for performance-based IADL measures, as these have been shown to be more sensitive in detecting subtle IADL impairments and predicting cognitive decline (Triebel et al., 2009; Puente et al., 2014; Sikkes and Rotrou, 2014). The aim of the current study was to assess whether participants' cognitive status, current depressive symptoms, and personality traits were differentially associated with IADL scores captured by subjective (self-report and informant-report) and objective (performance-based) measures, and whether this association captures variance above and beyond demographics and potential medical confounders.

MATERIALS AND METHODS

Participants

The present study reports data from participants in the Sydney Memory and Ageing Study (MAS), a longitudinal study of community-dwelling older adults aged 70–90 years that began in 2005 (Sachdev et al., 2010). Of the 8,914 individuals invited to participate, 1,037 participants were included in the baseline sample. Inclusion criteria were the ability to speak and write English sufficiently well to complete a psychometric assessment and self−report questionnaires. MAS baseline exclusion criteria were major psychiatric diagnoses, acute psychotic symptoms, or a current diagnosis of multiple sclerosis, motor neuron disease, developmental disability, progressive malignancy, or dementia. More detailed methods of recruitment and baseline demographics have been previously described by Sachdev et al. (2010).

Every 2 years, MAS participants undertook a detailed assessment with a trained research assistant during which they completed a comprehensive neuropsychological test battery, medical history, medical exam, and a series of questionnaires. Clinical diagnoses were made at each by an expert consensus panel who considered all available neuropsychological, clinical, and imaging data. At a 6-year follow-up, MCI was diagnosed using international consensus criteria (Winblad et al., 2004), and dementia was diagnosed according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (American Psychiatric Association, 1980), though participants meeting the later consensus criteria were excluded from the present analysis (see below).

For the present study, data from the 6-year follow-up are considered as this was the first time the performance-based assessment of IADL was administered. Of the 708 participants from the original MAS baseline sample included in the 6-year follow-up, there were 478 that completed the performance-based IADL measure. For the present study, further exclusion criteria were a consensus diagnosis of dementia at 6-year follow-up (n = 38), as such a diagnosis precluded administration of both the self-report and the performance-based IADL measures, and a non-English speaking background (i.e., not speaking English at a basic conversational level by the age of 9; n = 55), as cultural and linguistic variables, can bias standardized neuropsychological test results (Kochan et al., 2010). Thus, the final sample for this study comprised 385 participants, all of whom had an informant. Informants were friends or family members nominated by the participant who answered questions relating to the participant's memory, thinking, and daily functioning. Informants were required to have at least 1 h of contact with the participant per week, but reported, on average closer to 6 h (M = 5.89, SD = 8.12) of contact per week.

All participants and informants provided written informed consent to participate in this study, which was approved by the University of New South Wales Human Ethics Review Committee (HC:09382, 14327).

Measures

Self-Reported Instrumental Activities of Daily Living Assessment

To directly compare participants- and informant-reported IADL, we created a modified version of the Bayer-IADL. Where the original Bayer-IADL is designed to be administered to informants, we changed the pronoun of each question from third person to first person (e.g., "Does the *participant* have difficulty managing *his/her* medications?" to "Do *you* have difficulties managing *your* medications?" In this way, a direct comparison of participants' and informants' subjective appraisal of the same everyday activities could be captured while avoiding discrepancies between item phrasing or functional domains across measures. The modified self-report Bayer-IADL also takes on average 5 min to complete and is comprised of the same 25 items as the original Bayer-IADL (Hindmarch et al., 1998), which are also summed and divided by the number of items rated by the participant, with higher scores indicating more severe deficits.

Informant-Reported Instrumental Activities of Daily Living Assessment

The Bayer-Activities of Daily Living Scale (Bayer-IADL) (Hindmarch et al., 1998) was originally developed as an informant-based instrument to assess functional ability in the early stages of MCI and dementia. The Bayer-IADL typically takes 5 min to complete and is comprised of 25 items subdivided into three major areas: general ability to perform self-care and manage everyday activities (2 items), ability to perform specific everyday activities (18 items), and cognitive functions important for managing everyday life (5 items). Each item is introduced with the statement "Does the participant have difficulty. . ." and scores range from 1 (never) to 10 (always), with a "not applicable" or "unknown" option for each item. Scores are summed and divided by the number of items scored by the informant (i.e., not scored as "unknown" or "not applicable"), with higher scores indicating more severe deficits.

Performance-Based Instrumental Activities of Daily Living Assessment

The Sydney Test of Activities of daily living in Memory disorders (STAM) (Reppermund et al., 2017) is a performance-based measure of the ability to carry out a range of IADL. The STAM consists of nine items assessing the following domains of function: communication (making a phone call), dressing (putting on a shirt), handling finances (paying a bill by check), managing everyday activities (preparing the check for mailing), orientation to time (reading the time and setting an alarm), medication management (dispensing weekly medications), shopping (choosing items to make a simple recipe), counting money (calculating cost and counting money), and memory (recalling completed STAM activities). Each item on the STAM is scored on a 4-point scale, such that participants receive 1 point for each component completed correctly, for a maximum of 36 points where higher scores represent better performance. Each task also has a time limit, whereby participants were penalized if they went over time according to bracketed upper- and

lower- time cut-offs, with the STAM taking on average 15 min to complete. The complete scale, including instructions and item components for scoring, is provided as **Supplementary Material**.

Objective Cognitive Performance

Global cognition composite scores were based on participants' scores on a comprehensive neuropsychological test battery comprised of 10 tests that measured the domains of attention, language, executive function, visuospatial ability, memory, and verbal memory. Global cognition composite scores are presented as standardized z-scores and were derived as follows: Raw neuropsychological test scores were first converted to z-scores using the means and standard deviations (SDs) of a reference group comprised of 723 MAS participants from baseline that were classified as cognitively healthy (i.e., native English speakers with an MMSE score of 24 or above, no evidence of dementia or current depression, no history of delusions or hallucinations, and no major neurological disease, or significant head injuries). Composite domain scores were formed by averaging the z-scores of the component tests. Global cognition scores at each wave were calculated by averaging the domain scores. Global Cognition scores were standardized against the mean and SD (0 and 1, respectively) of the baseline reference group. More details about how cognitive domain and global cognition scores were calculated, and which tests comprised each cognitive domain, can be found in **Supplementary Material**.

Depression and Personality

The short form of the Geriatric Depression Scale (GDS) (Sheikh et al., 1986) was administered to assess participants' current depressive symptoms. The GDS is a self-reported measure that requires a yes/no response to 15 questions about current mood, with higher scores indicating more depressive symptoms. A score of 5 is the recommended cut-off for clinically relevant depression (Pocklington et al., 2016). We used the GDS version with item 9 as described in Brink (here item 12) (Brink, 1982). Personality traits were assessed using a modified 36-item version of the NEO Five-Factor Inventory NEO-FFI (Costa, 1992). The original NEO-FFI is a 60-item questionnaire that assesses the big five personality traits of Extraversion, Agreeableness, Openness, Neuroticism, and Conscientiousness. For the present study, only the latter three personality traits were considered. This decision was made to reduce participant burden at baseline and based on existing evidence suggesting Openness, Neuroticism, and Conscientiousness are most highly correlated with subjective cognitive complaints and incident dementia (Comijs et al., 2002; Duchek et al., 2007; Wilson et al., 2007). Twelve items relate to each of the three personality traits. Participants were asked to rate the degree to which they agree with each statement as it relates to their own beliefs or attributes on a 5-point scale, with higher scores indicating a higher prevalence of each personality trait.

Covariates

Covariates included basic demographics of age, sex, and education. Medical covariates of interest comprised self-reported arthritis and vision impairment, which we transformed into a composite variable given both impairments have been shown to significantly impact participants' ability to complete performance-based measures of IADL (Reppermund et al., 2017). Additionally, we included participants' self-reported number of medications (comprised of both prescription and over-the-counter medications), as a proxy for medical comorbidities, as well as the Framingham Risk Scores (D'Agostino et al., 2008) to capture cardiovascular disease (CVD) risk.

Statistical Analysis

Prior to analyses, all variables were screened for violations of the assumptions associated with univariate and multivariate tests. Both STAM and informant-reported Bayer-IADL scores were non-normally distributed; to avoid potentially inflating α, these scores were transformed (LOG_{10}), to improve normality and linearity, and univariate outliers were Windsorized. As analyses for both the raw and the transformed STAM and informant-report Bayer-IADL variables produced equivalent results, the raw data are listed in the table for ease of interpretation, but the transformed variables were used in analyses. Correlational analyses evaluated the relationship between participant-reported Bayer-IADL, informant-reported Bayer-IADL, and performance-based STAM scores.

Three hierarchical linear regressions were run to determine the predictive ability of participants' depression and personality traits on self-report, informant-report, and performance-based IADL scores, over and above the variance assumed by participant demographics, and medical covariates, and global cognitive function. For each of the three IADL outcome measures, a three-step hierarchical regression model was run in the same order. Step 1 always included participants' demographics (age, sex, and education) and medical comorbidities (vision impairment *and* arthritis, CVD risk score, total number of medications); Step 2 additionally included participants' Global Cognition composite score; and Step 3 additionally included participants' depression (GDS) and personality (NEO-FFI; neuroticism, conscientiousness, openness) scores. We ran additional *post-hoc* analyses to determine whether diagnostic status (i.e., NC vs. MCI) impacted the pattern of results. To do this, we ran the same series of hierarchical regressions, using the same steps and covariates outlined above, separately for the two groups.

All independent variables were assessed for multicollinearity with acceptable VIFs <2. Findings with a two-tailed $p < 0.05$ were considered statistically significant; analyses were performed using IBM SPSS Statistics 26 for Windows.

RESULTS

Participants

Table 1 presents participant characteristics for all test variables for the total sample (n = 385), and stratified by diagnostic group (i.e., NC vs. MCI). On average, participants were approximately 83 years old, were more often female (56%), and had just over 12 years of education. There was a low rate of depressive symptoms in the total sample, and participants self-reported more conscientiousness and openness compared to neuroticism. Group-level differences in participant characteristics for those

with NC vs. MCI emerged for several test variables. Compared with the NC participants, those with MCI were older and had higher CVD risk scores, significantly lower global cognition scores, and openness scores. In terms of outcome variables, participants with MCI had significantly higher (worse) self- and informant-reported IADL and significantly lower (worse) performance-based IADL scores.

Correlations Between Instrumental Activities of Daily Living Measures

Self-reported IADL scores were positively correlated with informant-reported scores, such that participants who reported more IADL difficulty also had informants who reported more IADL difficulty (r = 0.192, p < 0.001, N = 385). Conversely, performance-based IADL scores were significantly, and negatively, correlated with both self- (r = –0.151, p = 0.003, N = 385) and informant-reported IADL (r = –0.294, p < 0.001, N = 385). However, it is important to note that higher scores on the Bayer-IADL are indicative of worse IADL function whereas higher scores on the STAM indicate better IADL function. In sum, the strongest correlation to emerge was between informant-reported Bayer-IADL and STAM scores. Next, informant-reported, and participant-reported, Bayer-IADL scores emerged as significantly, but weakly, correlated (r < 2), as were participant-reported Bayer-IADL and STAM scores, which showed the weakest relationship among the three dependent variables.

Predictors of Self-Report Bayer-Instrumental Activities of Daily Living

Table 2 presents the results of a three-step hierarchical multiple regression predicting self-report IADL as the dependent variable. At step one, demographic and medical covariates contributed significantly to the regression model, $F(6, 348) = 2.24$, $p = 0.039$, and accounted for 3.7% of the variance in self-report IADL scores. Introducing the cognitive variables at step two contributed 2.8% of additional variance to the model, which was statistically significant, $F(1, 347) = 10.43$, $p = 0.001$. In step three, the addition of the depression and personality variables explained an additional 22.4% of the variance in self-report IADL, which was highly significant, $F(4, 343) = 27.06$, $p < 0.001$. Specifically, higher GDS and neuroticism scores, and lower conscientiousness scores, were related to higher (i.e., worse) self-reported IADL scores, suggesting participants' appraisals of their own functional ability were influenced by depressive symptoms and certain personality traits, over and above demographics, medical covariates, and global cognition.

Predictors of Informant-Report Bayer-Instrumental Activities of Daily Living

Table 3 presents the results of a three-step hierarchical multiple regression predicting informant-report IADL as the dependent variable. Participants' demographics and medical covariates contributed significantly to the model at step one, $F(6, 348) = 5.87$, $p < 0.001$, and accounted for 9.2% of the variance in informant-reported IADL. The addition of global cognition in step two explained an additional 4.7% of the variance in scores, which was highly significant, $F(1, 347) = 18.85$, $p < 0.001$. In step three, participants' depression and personality scores accounted for an additional 7.5% of the total model variance, over and above demographic, medical, and cognitive variables. This additional variance was statistically significant, $F(4, 343) = 8.21$, $p < 0.001$, and largely driven by participants' GDS scores, which uniquely explained 6% of the variation in informant-report IADL

TABLE 1 | Characteristics of the sample by total sample (All), normal cognition (NC), and Mild Cognitive Impairment (MCI).

	ALL (N = 385)	NC (N = 235)	MCI (N = 150)	Test statistic	P-value
Predictor variables					
Age, mean (SD)	83.21 (4.30)	82.81 (4.15)	83.83 (4.47)	t = –2.29	**0.023**
Sex—Female, n (%)	217 (56.4)	137 (58.3)	80 (53.3)	χ^2 = 0.338	0.345
Education (years), mean (SD)	12.05 (3.57)	12.24 (3.58)	11.74 (3.56)	t = 1.34	0.180
Arthritis + vision impaired, n (%)	31 (8.1)	15 (6.4)	16 (10.7)	χ^2 = 0.132	0.178
CVD risk score, mean (SD)	16.46 (3.45)	16.14 (6.78)	16.98 (3.23)	t = –2.29	**0.023**
Total no. medications mean (SD)	6.85 (3.36)	6.78 (3.20)	6.97 (3.60)	t = –0.54	0.592
†Global Cognition, mean (SD)	–0.19 (1.08)	0.33 (0.80)	–1.00 (0.95)	t = 14.72	**<0.001**
GDS, mean (SD)	1.43 (0.72)	1.37 (0.71)	1.52 (0.72)	t = –1.96	0.051
Neuroticism, mean (SD)	14.52 (6.67)	14.05 (6.53)	15.26 (6.85)	t = –1.72	0.087
Openness, mean (SD)	27.55 (5.88)	28.38 (5.62)	26.25 (6.06)	t = 3.48	**0.001**
Conscientiousness, mean (SD)	34.14 (5.93)	33.97 (5.87)	34.41 (6.03)	t = –0.71	0.478
Outcome variables					
Participant Bayer-IADL, mean (SD)	1.52 (0.62)	1.47 (0.51)	1.60 (0.67)	t = –2.22	**0.027**
Informant Bayer-IADL, mean (SD)	1.75 (0.83)	1.61 (0.72)	1.95 (0.87)	t = –4.13	**<0.001**
STAM score, mean (SD)	30.32 (4.59)	31.78 (3.59)	28.01 (5.15)	t = 8.46	**<0.001**

SD, standard deviation; CVD, cardiovascular disease; GDS, Geriatric Depression Scale; STAM, The Sydney Test of Activities of Daily Living in Memory Disorders.
†*Composite z-score. Bold values significant at p < 0.05.*

scores. These results suggest that while informants' appraisals of participants' functional ability are influenced by some variables expected to impact functional ability (e.g., years of education, number of medications, and global cognition scores), they are also significantly impacted by dynamic participant characteristics like depression.

Predictors of Performance-Based Instrumental Activities of Daily Living

Table 4 presents the results of a three-step hierarchical multiple regression predicting performance-based (STAM) IADL scores. As before, step one included participant demographics and medical covariates, which contributed significantly to the regression model, $F(6, 348) = 17.29$, $p < 0.001$, and accounted for 23.7% of the variance in performance-based IADL scores. Introducing global cognition at step two explained an additional 23% of the variation in performance-based IADL and this change in R^2 was highly significant, $F(1, 347) = 148.36$, $p < 0.001$, with global cognition scores alone explaining an additional 12% of the variance, over and above age, sex, and education and medical covariates. Interestingly, the addition of depression and personality in the final step did not contribute significantly to the final regression model, $F(4, 343) = 0.28$, $p = 0.508$. Together, these results suggest that scores on performance-based measures of IADLs are influenced most by participant characteristics like age, sex, education, and cognitive variables, as opposed to current depression or certain personality traits.

Post-hoc Analyses of Diagnostic Group Differences

Finally, to explore whether patterns of association differed by diagnostic status (i.e., NC vs. MCI), we re-ran our three hierarchical regressions, using the same steps and covariates as before, separately for the two groups. These results are presented in **Supplementary Material**. In general, the pattern of results was similar across outcome measures. **Supplementary Table 1** presents the results for the self-reported Bayer-IADL. In the fully adjusted model, GDS (NC: β = 0.248, $p < 0.001$; MCI: β = 0.392, $p < 0.001$) and conscientiousness (NC: β = –0.180, $p = 0.008$; MCI: β = –0.174, $p = 0.040$) remained significant for both groups. However, the significant effect of global cognition (β = –0.153, $p = 0.042$) and neuroticism (β = 0.241, $p = 0.001$) appear to be driven mostly by NC participants. **Supplementary Table 2** presents the results for the informant-reported Bayer-IADL. In the fully adjusted model, the effect of total number of medications was driven by the NC group (β = 0.185, $p = 0.008$) where the

TABLE 2 | Hierarchical linear regression predicting self-report Bayer-IADL.

Variable	β	*t*	sr^2	*R*	R^2	ΔR^2
Step 1				0.193	0.037	0.037*
Age	0.072	1.349	0.005			
Sex	–0.075	1.307	0.005			
Education	–0.008	0.158	0.000			
Arthritis + vision impairment	0.140	2.588*	0.019			
CVD risk	0.075	1.323	0.005			
Total no. medications	– 0.055	1.007	0.003			
Step 2				0.256	0.065	0.028**
Age	0.001	0.015	0.000			
Sex	–0.057	1.002	0.003			
Education	0.060	1.050	0.003			
Arthritis + vision impairment	0.136	2.558*	0.018			
CVD risk	0.072	1.292	0.004			
Total no. medications	–0.073	1.353	0.005			
Global cognition	–0.196	3.229**	0.028			
Step 3				0.538	0.290	0.224***
Age	–0.042	0.822	0.001			
Sex	–0.036	0.713	0.001			
Education	0.076	1.430	0.004			
Arthritis + vision impairment	0.062	1.310	0.004			
CVD risk	0.072	1.469	0.004			
Total no. medications	–0.111	2.331*	0.011			
Global cognition	–0.141	2.576*	0.014			
GDS	0.317	6.147***	0.078			
Neuroticism	0.166	3.063**	0.019			
Openness	0.047	0.896	0.002			
Conscientiousness	–0.185	3.639***	0.027			

*N = 403; *p < 0.05, **p < 0.01, ***p < 0.001.*
GDS, Geriatric Depression Scale (15-item version); CVD, cardiovascular disease.

effects of global cognition (β = –0.177, p = 0.042) and GDS (β = 0.501, $p < 0.001$) were driven by the MCI group. Openness also emerged as significant for MCI participants (β = 0.222, p = 0.010), where this was not significant in the original model. Finally, **Supplementary Table 3** presents the results for the performance-based IADL model. In the fully adjusted model, global cognition remained a significant predictor of total score for both groups (NC: β = 0.413, $p < 0.001$; MCI: β = 0.583, $p < 0.001$), though the association with education appears to be driven by the NC group, only (β = 0.215, p = 0.001).

DISCUSSION

This study examined the associations between measures of cognitive function, depressive symptoms, and personality traits and different methods of assessing IADL in a community-dwelling older sample, controlling for demographics and potential medical confounders. When considering the total sample, we found that self-reported worse IADL was associated with worse global cognition scores, depressive symptoms, higher scores on neuroticism, and lower scores on conscientiousness. Worse informant-reported IADL function was also associated with worse global cognition scores and increased depression scores but was not associated with participants' personality traits. Finally, performance-based IADL scores were associated with age, sex, education, and global cognition, but were not associated with depression or personality traits in any significant way. In addition, the three measures of IADL were only weakly correlated, which may reflect how each measure is variably influenced by demographics, mood, personality traits, and cognition. *Post-hoc* analyses considering diagnostic groups revealed that in some instances associations with the full sample analysis held (e.g., GDS and conscientiousness predicting participant-reported IADL and global cognition predicting performance-based IADL) in other cases differences were driven by one group over the other (e.g., GDS and global cognition predicting informant-reported IADL for MCI only). However, given we were most interested in understanding how these variables across the predementia spectrum influence different measures of functional ability, we focus our attention on the group level analyses.

Our findings that self-reported IADL impairment is significantly associated with higher neuroticism and lower conscientiousness align with other previous reports (Chapman et al., 2007; Suchy et al., 2010; Puente et al., 2015). Given

TABLE 3 | Hierarchical linear regression predicting informant-report Bayer-IADL.

Variable	β	*T*	sr^2	*R*	R^2	ΔR^2
Step 1				0.303	0.092	0.092***
Age	0.124	2.400*	0.015			
Sex	–0.127	2.269*	0.013			
Education	–0.131	2.514*	0.016			
Arthritis + vision impairment	0.084	1.598	0.007			
CVD risk	0.045	0.819	0.002			
Total no. medications	0.170	3.238**	0.027			
Step 2				0.372	0.139	0.047***
Age	0.033	0.595	0.001			
Sex	–0.104	1.890	0.009			
Education	–0.043	0.781	0.002			
Arthritis + vision impairment	0.079	1.552	0.006			
CVD risk	0.042	0.774	0.001			
Total no. medications	0.147	2.847**	0.020			
Global cognition	–0.253	4.341***	0.047			
Step 3				0.463	0.214	0.075***
Age	–0.012	0.224	0.000			
Sex	–0.104	1.942	0.009			
Education	–0.065	1.161	0.003			
Arthritis + vision impairment	0.039	0.777	0.001			
CVD risk	0.027	0.520	0.001			
Total no. medications	0.115	2.296*	0.012			
Global cognition	–0.232	4.040***	0.037			
GDS	0.280	5.161***	0.061			
Neuroticism	0.018	0.311	0.000			
Openness	0.095	1.710	0.007			
Conscientiousness	–0.007	0.127	0.000			

*N = 355; *p < 0.05, **p < 0.01, ***p < 0.001.*
GDS, Geriatric Depression Scale (15-item version); CVD, cardiovascular disease.

TABLE 4 | Hierarchical linear regression predicting performance-based IADL.

Variable	β	*t*	sr^2	*R*	R^2	ΔR^2
Step 1				0.479	0.230	0.230***
Age	–0.304	6.388***	0.090			
Sex	0.165	3.214**	0.023			
Education	0.342	7.136***	0.113			
Arthritis + vision impairment	–0.025	0.520	0.001			
CVD risk	0.017	0.332	0.000			
Total no. medications	–0.092	1.905	0.008			
Step 2				0.679	0.460	0.231***
Age	–0.101	2.333*	0.008			
Sex	0.114	2.622**	0.011			
Education	0.146	3.381**	0.018			
Arthritis + vision impairment	–0.015	0.378	0.000			
CVD risk	0.025	0.582	0.001			
Total no. medications	–0.040	0.986	0.002			
Global cognition	0.562	12.180***	0.231			
Step 3				6.82	0.466	0.005
Age	–0.086	1.932	0.006			
Sex	0.104	2.357*	0.009			
Education	0.145	3.138**	0.015			
Arthritis + vision impairment	–0.009	0.220	0.000			
CVD risk	0.028	0.648	0.001			
Total no. medications	–0.043	1.041	0.002			
Global cognition	0.561	11.839***	0.218			
GDS	–0.026	0.589	0.001			
Neuroticism	0.065	1.386	0.003			
Openness	0.022	0.474	0.000			
Conscientiousness	0.066	1.498	0.003			

*N = 355; *p < 0.05, **p < 0.01, ***p < 0.001.*
GDS, Geriatric Depression Scale (15-item version); CVD, cardiovascular disease.

that these personality traits are not significantly associated with informant-reported and performance-based IADL, these findings may imply that individuals with such traits are more prone to over-report IADL impairment. Indeed, individuals who score higher on measures of neuroticism tend to experience more, and ruminate over, negative emotions like stress and anxiety (Widiger and Oltmanns, 2017). As such, individuals with higher neuroticism may tend to over-report IADL impairment due to emotional lability, pessimistic views of oneself, and feelings of vulnerability (Widiger and Oltmanns, 2017). In addition, conscientiousness is characterized by being diligent, organized, self-disciplined, and determined. Previous studies have suggested that individuals with higher conscientiousness may either experience less functional impairment (due to better able to plan and execute tasks),Krueger et al. (2006) or be less willing to admit ADL impairments (Roy et al., 2016; Williams et al., 2017). Our findings indicate that the reverse may also be true, where individuals with low conscientiousness may be less reluctant to reveal IADL impairments and maybe even more prone to over-report impairment. These findings suggest that self-reported IADL may be biased by such personality traits and thus, may not truly reflect the individual's level of function (Suchy et al., 2010). This has important clinical implications when considering the validity of a patient's self-reported function.

Our findings further revealed that depression was significantly associated with worse subjective IADL ratings reported by both the participant and their informant, suggesting that individuals with depressive symptoms, and their informants, may tend to overreport IADL impairments. These results support previous research showing depressive symptoms predict worse self-reported IADL performance over time (Ryu et al., 2016; Sutin et al., 2016). However, the direction of this relationship remains unclear. That is, previous research has suggested that depression may precede functional impairment (Kong et al., 2019) as depressive symptoms are highly correlated with loss of energy and motivation, decreased activity, poorer health behaviors, and psychomotor slowing (Lenze et al., 2001; Schillerstrom et al., 2008). On the other hand, declines in functional ability may be the precursor to depression (Schillerstrom et al., 2008) as a loss of independence for daily tasks is highly correlated with lower perceived quality of life and worse self-reported life satisfaction (Meltzer et al., 2012). Only longitudinal studies, however, can determine the impact of depression and its comorbidities on functional decline.

The finding that depressive symptoms were not associated with IADL function as measured by an objective performance-based measure favors the hypothesis that depressive symptoms may be related to overreporting of IADL impairment rather than actual functional decline. One explanation for this relationship

is that participants experiencing depressive symptoms may more easily recollect negative instances when they could not perform daily tasks. This is supported by research on the relationship between mood and memory (Williams, 1999). However, it is important to note that participants in our study did not have clinical depression and that indeed many older adults experience depressive symptoms (Dozeman et al., 2010). Thus, to disentangle the complex relationship between increased depression and worse subjectively reported IADL and preserved performance-based objective IADL, a prospective examination of these associations using longitudinal data is required. Nevertheless, our results confirm the importance of screening for depressive symptoms when assessing IADL.

Interestingly, all three measures of IADL were significantly associated with global cognition. However, this association was strongest for objective performance-based IADL ability in comparison to subjectively reported IADL. Nearly 22% of the total variance in performance-based IADL scores was predicted by global cognition compared to less than 2% for self-report and less than 4% for informant-report. Again, this may be due to the impact of depressive symptoms and personality traits on subjective impressions of IADL ability. Taken together, our results suggest that performance-based measures are more reflective of the individual's cognitive status compared to performance-based measures. This is consistent with previous research which has demonstrated that performance-based IADL measures may be more sensitive in discriminating between cognitive status (Pereira et al., 2010).

Strengths of our study are the relatively large and well-characterized sample, expert consensus clinical diagnosis of NC vs. MCI, the inclusion of important medical covariates, and the three types of IADL measures. To our knowledge, subjective self- and informant-reported IADL and performance-based IADL have not been examined together with depression and personality in one cohort. This study has certain limitations. Firstly, the data analyzed are cross-sectional and provide limited information on whether depressive symptoms are the antecedent or consequences of functional impairment in IADL. Further longitudinal studies should endeavor to explore this issue, particularly relating to performance-based IADL measures. We do not include measures of informant-reported depression, personality, or perceived burden, which may also impact their subjective reports of participants' IADL ability. A final limitation is that this study only examined individuals with normal cognition or MCI. Thus, the findings do not extend to individuals with dementia. As dementia can have a long prodromal period, future research should look at whether the relationships reported here are similar or different for persons with mild or moderate dementia diagnoses.

In sum, while self-reported IADL function was significantly associated with the participant's current level of global cognition it was more strongly associated with depressive symptoms and personality traits. Informant-reported IADL was also influenced by depressive symptoms but was not impacted by personality traits and was strongly associated with global cognition. Finally, performance-based IADL scores were not significantly associated with the participant's depression and personality traits; instead, they were mostly accounted for by the participant's age, sex, education, and global cognition—all variables known to be associated with functional ability in older adults. We argue that performance-based IADL measures are preferable when trying to accurately assess everyday functional ability and its relationship to cognitive status. Where performance-based measures are not available (e.g., in some clinical settings), informant ratings should be sought as they are less influenced by the participant's personality and mood compared to self-reports.

DATA AVAILABILITY STATEMENT

The terms of consent for research participation stipulate that an individual's data can only be shared outside of the MAS investigators group if the group has reviewed and approved the proposed secondary use of the data. This consent applies regardless of whether data have been de-identified. Access is mediated via a standardised request process managed by the CHeBA Research Bank, who can be contacted at ChebaData@unsw.edu.au, or *via* KN, k.numbers@unsw.edu.au.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the University of New South Wales Human Research Ethics Committee. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

KN designed this study, wrote this manuscript, completed all analyses, and prepared this manuscript for submission. SJ, HB, PS, and BD provided statistical, neuropsychological, and medical guidance, and reviewed this manuscript and revisions. SR provided detailed feedback and guidance at each step of design and manuscript drafting as senior author. All authors contributed to the article and approved the submitted version.

FUNDING

National Health and Medical Research Council (NHMRC) of Australia Program Grants, Grant/Award Numbers: ID350833, ID568969, and APP1093083.

ACKNOWLEDGMENTS

We would like to thank Nicole A. Kochan for his assistance with neuropsychological guidance for the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnagi.2022.829544/full#supplementary-material

REFERENCES

Albert, S. M., Michaels, K., Padilla, M., Pelton, G., Bell, K., Marder, K., et al. (1999). Functional significance of mild cognitive impairment in elderly patients without a dementia diagnosis. *Am. J. Geriat. Psychiatry* 7, 213–220.

American Psychiatric Association (1980). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association.

Argüelles, S., Loewenstein, D. A., Eisdorfer, C., and Argüelles, T. (2001). Caregivers' judgments of the functional abilities of the Alzheimer's disease patient: impact of caregivers' depression and perceived burden. *J. Geriat. Psychiat. Neurol.* 14, 91–98. doi: 10.1177/089198870101400209

Brink, T. L. (1982). Geriatric depression and hypochondriasis: incidence, interaction, assessment and treatment. *Psychotherapy* 19, 506–511.

Chapman, B., Duberstein, P., and Lyness, J. M. (2007). Personality traits, education, and health-related quality of life among older adult primary care patients. *J. Gerontol. Series B* 62, 343–352.

Coman, L., and Richardson, J. (2006). Relationship between self-report and performance measures of function: a systematic review. *Can. J. Aging* 25, 253–270.

Comijs, H., Deeg, D., Dik, M., Twisk, J., and Jonker, C. J. (2002). Memory complaints, the association with psycho-affective and health problems and the role of personality characteristics: a 6-year follow-up study. *J. Affect. Disord.* 72, 157–165. doi: 10.1016/s0165-0327(01)00453-0

Costa, T. Jr. (1992). The NEO-PI-R professional manual: revised NEO Five-Factor Inventory.(NEO-FFI). *Psychol. Assess. Resour.* 1992:101.

D'Agostino, R. B. Sr., Vasan, R. S., Pencina, M. J., Wolf, P. A., Cobain, M., Massaro, J. M., et al. (2008). General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation* 117, 743–753.

de Paula, J. J., Diniz, B. S., Bicalho, M. A., Albuquerque, M. R., Nicolato, R., de Moraes, E. N., et al. (2015). Specific cognitive functions and depressive symptoms as predictors of activities of daily living in older adults with heterogeneous cognitive backgrounds. *Front. Aging Neurosci.* 2015:7. doi: 10.3389/fnagi.2015.00139

Dozeman, E., van Marwijk, H. W., van Schaik, D. J., Stek, M. L., van der Horst, H. E., Beekman, A. T., et al. (2010). High incidence of clinically relevant depressive symptoms in vulnerable persons of 75 years or older living in the community. *Aging Ment. Health* 14, 828–833.

Duchek, J. M., Balota, D. A., Storandt, M., and Larsen, R. J. T. (2007). The power of personality in discriminating between healthy aging and early-stage Alzheimer's disease. *J. Gerontol. B Psychol. Sci. Soc. Sci.* 62, 353–361.

Gold, D. A. (2012). An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. *J. Clin. Exp. Neuropsychol.* 34, 11–34.

Goldberg, T. E., Koppel, J., Keehlisen, L., Christen, E., Dreses-Werringloer, U., Conejero-Goldberg, C., et al. (2010). Performance-based measures of everyday function in mild cognitive impairment. *Am. J. Psychiatry.* 167:845.

Griffith, H., Belue, K., Sicola, A., Krzywanski, S., Zamrini, E., Harrell, L., et al. (2003). Impaired financial abilities in mild cognitive impairment: a direct assessment approach. *Neurology* 60, 449–457.

Hindmarch, I., Lehfeld, H., de Jongh, P., and Erzigkeit, H. (1998). The Bayer activities of daily living scale (B-ADL). *Dement. Geriat. Cogn. Dis.* 9, 20–26.

Karakurt, P., and Unsal, A. (2013). Fatigue, anxiety and depression levels, activities of daily living of patients with chronic obstructive pulmonary disease. *Int. J. Nurs. Pract.* 19, 221–231.

Kempen, G. I., Jelicic, M., and Ormel, J. (1997). Personality, chronic medical morbidity, and health-related quality of life among older persons. *Health Psychol.* 16:539.

Kochan, N. A., Slavin, M. J., Brodaty, H., Crawford, J. D., Trollor, J. N., Draper, B., et al. (2010). Effect of different impairment criteria on prevalence of "objective" mild cognitive impairment in a community sample. *Am. J. Geriat. Psychiat.* 18, 711–722. doi: 10.1097/jgp.0b013e3181d6b6a9

Kong, D., Solomon, P., and Dong, X. (2019). Depressive Symptoms and Onset of Functional Disability Over 2 Years: a Prospective Cohort Study. *J. Am. Geriatr. Soc.* 67, S538–S544. doi: 10.1111/jgs.15801

Krueger, K. R., Wilson, R. S., Shah, R. C., Tang, Y., and Bennett, D. A. (2006). Personality and incident disability in older persons. *Age Ageing* 35, 428–433.

Lenze, E. J., Rogers, J. C., Martire, L. M., Mulsant, B. H., Rollman, B. L., Dew, M. A., et al. (2001). The association of late-life depression and anxiety with physical disability: a review of the literature and prospectus for future research. *Am. J. Geriat. Psychiatry* 9, 113–135.

Martyr, A., Nelis, S. M., and Clare, L. (2014). Predictors of perceived functional ability in early-stage dementia: self-ratings, informant ratings and discrepancy scores. *Internat. J. Geriat. Psychiatry* 29, 852–862. doi: 10.1002/gps.4071

Meltzer, H., Bebbington, P., Brugha, T., McManus, S., Rai, D., Dennis, M. S., et al. (2012). Physical ill health, disability, dependence and depression: Results from the 2007 national survey of psychiatric morbidity among adults in England. *Disabil. Health J.* 5, 102–110. doi: 10.1016/j.dhjo.2012.02.001

Mitchell, M., and Miller, L. S. (2008). Executive functioning and observed versus self-reported measures of functional ability. *Clin. Neuropsychol.* 22, 471–479. doi: 10.1080/13854040701336436

Mlinac, M. E., and Feng, M. C. (2016). Assessment of activities of daily living, self-care, and independence. *Archiv. Clin. Neuropsychol.* 31, 506–516.

Moore, D. J., Palmer, B. W., Patterson, T. L., and Jeste, D. V. (2007). A review of performance-based measures of functional living skills. *J. Psychiatric Res.* 41, 97–118.

Pereira, F. S., Yassuda, M. S., Oliveira, A. M., Diniz, B. S., Radanovic, M., Talib, L. L., et al. (2010). Profiles of functional deficits in mild cognitive impairment and dementia: benefits from objective measurement. *J. Internat. Neuropsychol. Soc.* 16, 297–305. doi: 10.1017/S1355617709991330

Pfeifer, L., Drobetz, R., Fankhauser, S., Mortby, M. E., Maercker, A., Forstmeier, S., et al. (2013). Caregiver rating bias in mild cognitive impairment and mild Alzheimer's disease: impact of caregiver burden and depression on dyadic rating discrepancy across domains. *Internat. Psychogeriat.* 25, 1345–1355.

Pocklington, C., Gilbody, S., Manea, L., and McMillan, D. (2016). The diagnostic accuracy of brief versions of the Geriatric Depression Scale: a systematic review and meta-analysis. *Int. J. Geriatr. Psychiatry* 31, 837–857. doi: 10.1002/gps.4407

Puente, A. N., Lindbergh, C. A., and Miller, L. S. (2015). Personality's association with IADLs in community dwelling older adults. *Internat. J. Geriat. Psychiatry* 30, 950–956.

Puente, A. N., Terry, D. P., Faraco, C. C., Brown, C. L., and Miller, L. S. (2014). Functional impairment in mild cognitive impairment evidenced using performance-based measurement. *J. Geriat. Psychiat. Neurol.* 27, 253–258.

Reppermund, S., Birch, R. C., Crawford, J. D., Wesson, J., Draper, B., Kochan, N. A., et al. (2017). Performance-based assessment of instrumental activities of daily living: validation of the Sydney Test of Activities of Daily Living in Memory Disorders (STAM). *J. Am. Med. Direct. Assoc.* 18, 117–122. doi: 10.1016/j.jamda.2016.08.007

Roy, S., Ficarro, S., Duberstein, P., Chapman, B. P., Dubovsky, S., Paroski, M., et al. (2016). Executive function and personality predict instrumental activities of daily living in Alzheimer disease. *Am. J. Geriat. Psychiatry* 24, 1074–1083.

Ryu, S. Y., Lee, S. B., Kim, T. W., and Lee, T. J. (2016). Subjective memory complaints, depressive symptoms and instrumental activities of daily living in mild cognitive impairment. *Internat. Psychogeriatr.* 28, 487–494.

Sachdev, P. S., Brodaty, H., Reppermund, S., Kochan, N. A., Trollor, J. N., Draper, B., et al. (2010). The Sydney Memory and Ageing Study (MAS): methodology and baseline medical and neuropsychiatric characteristics of an elderly epidemiological non-demented cohort of Australians aged 70–90 years. *Int. Psychogeriatr.* 22, 1248–1264. doi: 10.1017/S1041610210001067

Schillerstrom, J. E., Royall, D. R., and Palmer, R. F. (2008). Depression, disability and intermediate pathways: a review of longitudinal studies in elders. *J. Geriatric Psychiatry Neurol.* 21, 183–197. doi: 10.1177/0891988708320971

Schmitter-Edgecombe, M., Parsey, C., and Cook, D. J. (2011). Cognitive correlates of functional performance in older adults: comparison of self-report, direct observation, and performance-based measures. *J. Internat. Neuropsychol Soc.* 17, 853–864.

Sheikh, J. I., Yesavage, J. A., and Health, M. (1986). Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin. Gerontolog.* 5, 165–173. doi: 10.3109/09638288.2010.503835

Sikkes, S. A., and Rotrou, J. D. (2014). A qualitative review of instrumental activities of daily living in dementia: what's cooking? *Neurodegen. Dis. Manag.* 4, 393–400. doi: 10.2217/nmt.14.24

Steward, K. A., Kennedy, R., Erus, G., Nasrallah, I. M., and Wadley, V. G. (2019). Poor awareness of IADL deficits is associated with reduced regional

brain volume in older adults with cognitive impairment. *Neuropsychologia* 129, 372–378. doi: 10.1016/j.neuropsychologia.2019.04.023

Storeng, S. H., Sund, E. R., and Krokstad, S. (2018). Factors associated with basic and instrumental activities of daily living in elderly participants of a population-based survey: the Nord-Trondelag Health Study, Norway. *BMJ Open* 2018:8. doi: 10.1136/bmjopen-2017-018942

Suchy, Y., Williams, P. G., Kraybill, M. L., Franchow, E., and Butner, J. (2010). Instrumental Activities of Daily Living Among Community-Dwelling Older Adults: personality Associations With Self-Report, Performance, and Awareness of Functional Difficulties. *J. Gerontol.* 65B, 542–550. doi: 10.1093/geronb/gbq037

Sutin, A. R., Stephan, R., Luchetti, R., and Artese, A. L. (2016). The five-factor model of personality and physical inactivity: a meta-analysis of 16 samples. *J. Res. Person.* 63, 22–28.

Tabert, H. M., Albert, S. M., Borukhova-Milov, L., Camacho, Y., Pelton, G., Liu, X., et al. (2002). Functional deficits in patients with mild cognitive impairment: prediction of AD. *Neurology* 58, 758–764.

Taş, U., Verhagen, A. P., Bierma-Zeinstra, S. M., Hofman, A., Odding, E., Pols, H. A., et al. (2007). Incidence and risk factors of disability in the elderly: the Rotterdam Study. *Prev. Med.* 44:272.

Teng, E., Becker, B. W., Woo, E., Knopman, D. S., Cummings, J. L., Lu, P. H., et al. (2010). Utility of the Functional Activities Questionnaire for distinguishing mild cognitive impairment from very mild Alzheimer's disease. *Alzheimer Dis. Assoc. Dis.* 24:348.

Triebel, L. K., Martin, R., Griffith, H. R., Marceaux, J., Okonkwo, O. C., Harrell, L., et al. (2009). Declining financial capacity in mild cognitive impairment: a 1-year longitudinal study. *Neurology* 73, 928–934. doi: 10.1212/WNL.0b013e3181b87971

Votruba, K. L., Persad, C., and Giordani, B. (2015). Patient mood and instrumental activities of daily living in Alzheimer disease: relationship between patient and caregiver reports. *J. Geriat. Psychiat. Neurol.* 28, 203–209.

Widiger, T. A., and Oltmanns, J. R. (2017). Neuroticism is a fundamental domain of personality with enormous public health implications. *World Psychiatry* 16:144. doi: 10.1002/wps.20411

Williams, J. (1999). *10 Depression and the specificity of autobiographical memory. Remembering our past: Studies in autobiographical memory*. Cambridge, MA: Cambridge University Press. 1999:244.

Williams, P. G., Rau, H. K., Suchy, Y., Thorgusen, S. R., and Smith, T. W. (2017). On the validity of self-report assessment of cognitive abilities: Attentional control scale associations with cognitive performance, emotional adjustment, and personality. *Psycholog. Assess.* 29:519.

Wilson, R. S., Schneider, J. A., Arnold, S. E., Bienias, J. L., and Bennett, D. A. (2007). Conscientiousness and the Incidence of Alzheimer Disease and Mild Cognitive Impairment. *Archiv. Gen. Psychiat.* 64, 1204–1212.

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., et al. (2004). Mild cognitive impairment–beyond controversies, towards a consensus: report of the International Working Group on Mild Cognitive Impairment. *J. Intern. Med.* 256, 240–246. doi: 10.1111/j.1365-2796.2004.01380.x

Zanetti, O., Geroldi, C., Frisoni, G. B., Bianchetti, A., and Trabucchi, M. (1999). Contrasting results between caregiver's report and direct assessment of activities of daily living in patients affected by mild and very mild dementia: the contribution of the caregiver's personal characteristics. *J. Am. Geriat. Soc.* 47, 196–202. doi: 10.1111/j.1532-5415.1999.tb04578.x

Conflict of Interest: HB was advisory board member or consultant to Biogen, Nutricia, Roche and Skin2Synapse. He was Medical/Clinical Advisory Board member for Montefiore Homes and Cranbrook Care.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Aging Neuroscience

TYPE Original Research
PUBLISHED 27 July 2022
DOI 10.3389/fnagi.2022.912972

OPEN ACCESS

EDITED BY
Robert Fellows,
United States Department of Veterans Affairs, United States

REVIEWED BY
Hui Zhao,
Nanjing Drum Tower Hospital, China
Diogo Goulart Corrêa,
Federal University of Rio de Janeiro, Brazil

*CORRESPONDENCE
Yoshiaki Tamura
tamurayo@tmghig.jp

†PRESENT ADDRESS
Keigo Shimoji,
Department of Radiology, Juntendo University Graduate School of Medicine,
Tokyo, Japan

SPECIALTY SECTION
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 05 April 2022
ACCEPTED 30 June 2022
PUBLISHED 27 July 2022

CITATION
Tamura Y, Shimoji K, Ishikawa J, Murao Y, Yorikawa F, Kodera R, Oba K, Toyoshima K, Chiba Y, Tokumaru AM and Araki A (2022) Association between white matter alterations on diffusion tensor imaging and incidence of frailty in older adults with cardiometabolic diseases.
Front. Aging Neurosci. 14:912972.
doi: 10.3389/fnagi.2022.912972

Association between white matter alterations on diffusion tensor imaging and incidence of frailty in older adults with cardiometabolic diseases

Yoshiaki Tamura[1,2]*, Keigo Shimoji[3†], Joji Ishikawa[2,4], Yuji Murao[2], Fumino Yorikawa[2], Remi Kodera[1], Kazuhito Oba[1,2], Kenji Toyoshima[1,2], Yuko Chiba[1], Aya M. Tokumaru[3] and Atsushi Araki[1,2]

[1]Department of Diabetes, Metabolism, and Endocrinology, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan , [2]The Center for Comprehensive Care and Research for Prefrailty, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan, [3]Department of Diagnostic Radiology, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan , [4]Department of Cardiology, Tokyo Metropolitan Geriatric Hospital, Tokyo, Japan

Diffusion tensor imaging (DTI) can be used for the early detection of abnormal changes in the integrity of cerebral white matter tracts, and we have previously reported that these changes are associated with indices of early atherosclerotic lesions. Although these changes have been demonstrated to be associated with the incidence of frailty in older adults, no studies have investigated this relationship in patients at high risk for vascular disease. In this longitudinal study, we followed outpatients with cardiometabolic diseases for a maximum of 6 years (median, 3 years) and evaluated the association of baseline DTI data of seven white matter tracts with the incidence of frailty. The modified version of the Cardiovascular Health Study criteria and the Kihon Checklist were used as indices of frailty; fractional anisotropy (FA) and mean diffusivity (MD) were used as indices of white matter changes. Patients who developed frailty based on both indices had low FA and high MD in many of the tracts tested, with the most significant difference found in the MD of the anterior thalamic radiation (ATR). Cox proportional hazard model analysis revealed a significantly high risk of frailty defined by both indices in the groups with high MD values in the left ATR. Similar results were found in patients with diabetes mellitus but not in those without diabetes mellitus. Therefore, abnormalities in the integrity of the left ATR could be associated with the progression of frailty in older adults with cardiometabolic disease, particularly those with diabetes mellitus.

KEYWORDS

diabetes mellitus, diffusion tensor imaging, frailty, older adults, white matter alteration

Introduction

The population of developed countries is currently aging significantly, thereby escalating the economic burden of medical and nursing care for older adults. Frailty is a condition in which older people become vulnerable to various stresses outside the process of aging (Fried et al., 2001). It has been established that physically frail older adults often need nursing care and have an increased death rate (Ensrud et al., 2008).

To decrease the incidence of frailty, identifying high-risk individuals and early intervention are crucial. Hence, various potential frailty markers are currently under investigation. White matter hyperintensity (WMH) detected by fluid-attenuated inversion recovery magnetic resonance imaging (MRI) reflects small vessel disease in the cerebrum; we had previously reported that the WMH volumes in the whole cerebrum or periventricular lesions are associated with impairments in instrumental activities of daily living (Tamura et al., 2017). Diffusion tensor imaging (DTI), an imaging technology that uses diffusion-weighted images in multiple dimensions, has recently been developed. It uses the alteration of the diffusion anisotropy of water molecules, which in normal tracts move along the direction of the major axis of the neuron bundles. DTI has several advantages over WMH analysis, including that it can depict white matter lesions in tracts connecting various cerebral areas, such as the frontal cortex and thalamus (Assaf and Pasternak, 2008). Moreover, slight changes in white matter tracts can be detected earlier than with the classic WMH imaging method.

Recently, we showed that changes in DTI values in some white matter tracts are associated with indices of subclinical atherosclerosis, such as the ankle-brachial index (Tamura et al., 2021a). In a cross-sectional study of the same patients, we also demonstrated that abnormalities in white matter integrity in certain tracts, such as the left anterior thalamic radiation (ATR, lATR) and right inferior fronto-occipital fasciculus (IFOF, rIFOF), are associated with sarcopenia and its component factors (Tamura et al., 2021b). To the best of our knowledge, only one cross-sectional study (Avila-Funes et al., 2017) and one longitudinal study (Maltais et al., 2020) have been conducted on the association between white matter integrity and frailty. However, the participants of both studies were community-dwelling older adults; the influence of white matter lesions on frailty among those at risk for vascular diseases is unknown.

Accumulation of cardiometabolic diseases could be a risk of frailty. In a recent report, it has been shown that as the number of cardiometabolic diseases increases [hypertension, diabetes mellitus (DM), dyslipidemia, cardiac diseases, and stroke], the prevalence of frailty also increases in parallel (Gao et al., 2021). Among these diseases, DM is an independent strong risk factor for frailty. Since it has been shown in a meta-analysis that the incidence of frailty is significantly high in patients with diabetes mellitus and those who are affected with frailty are at high risk for mortality (Hanlon et al., 2020), it is of importance to screen for frailty and appropriately intervene in these high-risk patients at its early stage. In this context, it is meaningful to investigate the association between white matter alteration and the incidence of frailty in patients with cardiometabolic diseases, especially those with DM.

This study aimed to investigate the association between changes in DTI values of white matter tracts and the incidence of frailty, defined using the modified version of the Cardiovascular Health Study (CHS) criteria (mCHS) and Kihon Checklist (KCL), in outpatients with cardiometabolic diseases including those with DM (62.5%) and further verify whether these changes could be used as an early predictive marker of frailty in these high risk population.

Patients and methods

Patients

This longitudinal study analyzed data from 184 patients of a previous study (Tamura et al., 2021b). As described previously, they were outpatients aged ≥65 years at the frailty clinic and underwent brain MRI. In principle, these patients are recommended to visit the frailty clinic for an annual assessment of frailty status. Most patients were treated at the Departments of Diabetes and Cardiology and had cardiometabolic diseases, such as hypertension, diabetes mellitus, dyslipidemia, ischemic heart disease, and heart failure. Since patients with these backgrounds are at high risk for ischemic stroke, brain MRI was frequently performed to rule out stroke and to evaluate the ischemic status. The other inclusion criteria for the present analysis were: (1) data available on frailty status assessed using either of the two diagnostic criteria of frailty explained below, a modified version of the Cardiovascular Health Study (CHS) criteria (mCHS; Tamura et al., 2018) or Kihon Checklist (KCL), at baseline and at least one more time point till October 6, 2021; (2) the patients who were not frail at baseline (unaffected) by either criterion that met condition (1). Accordingly, data from 184 patients were eligible for inclusion. The number of patients who had both mCHS and KCL data at baseline was 182. For each of the mCHS and KCL, data from 137 and 154 patients unaffected at baseline were used, respectively. DM was diagnosed as previously described (Tamura et al., 2021b). The study protocol was approved by the ethics committee of the Tokyo Metropolitan Geriatric Hospital (R15-20, 19-03). Written informed consent was obtained from all patients. Patients with a history of brain neoplasm, neurosurgery, traumatic brain injury, psychiatric or neurological illness,

or gross intracranial hemorrhage were excluded from this study.

Assessment of frailty

Various diagnostic criteria for frailty have been proposed. The gold standard for diagnosing physical frailty is the CHS criteria, which is based on the abovementioned definition by Fried et al. (2001). Many modified versions of the CHS criteria have been developed, including the J-CHS, which is a modified version of the CHS criteria that is suitable for use in Japanese older adults (Satake et al., 2017a); the J-CHS was modified in 2020 (Satake and Arai, 2020). We have previously used the mCHS criteria that comprised almost the same items as the J-CHS and reported the prevalence of frailty in our outpatient clinic (Tamura et al., 2018).

Recently, the concept of frailty has been broadened to include comprehensive geriatric assessment, which includes the assessment of physical and cognitive function and social status. KCL defines criteria based on this broadened concept of frailty. High KCL scores have been associated with several CHS frailty phenotypes (Satake et al., 2016), mortality, disability (Satake et al., 2017b), and hospitalization (Koyama et al., 2022) in community-dwelling Japanese older adults.

In this study, the patients' frailty status was determined using the mCHS and KCL. Patients with $\geq$3 out of 5 points on the mCHS were considered frail. However, based on recent revisions in the J-CHS (Satake and Arai, 2020), we revised the cutoff value of grip strength in men from 26 kg to 28 kg. The KCL is a multidomain questionnaire for evaluating frailty status, and patients with a score $\geq$8 were classified as frail (Satake et al., 2016). The Mini-Mental State Examination (MMSE) and the Japanese version of the Montreal Cognitive Assessment (MoCA-J) were used to evaluate cognitive function, and those with MMSE $\geq$ 24 and MoCA-J $\leq$ 25 were defined as suspected mild cognitive impairment (MCI) cases. The International Physical Activity Questionnaire was used to calculate the patients' physical activity (metabolic equivalents $\times$ min/week), as described by Craig et al. (2003).

MRI acquisition and DTI analysis

As performed in our previous study (Tamura et al., 2021b), MRI acquisition was carried out using a Discovery MR750w 3.0 MRI system with a 16-channel head coil (General Electric Healthcare, Milwaukee, WI, USA).

DTI acquisitions were executed using a single-shot spin-echo echoplanar imaging sequence with diffusion gradient encoding schemes consisting of 32 non-collinear directions with a *b*-value of 1,000 s/mm^2 and one non-diffusion-weighted image with a *b*-value of 0 s/mm^2, in the axial plane along the anterior-posterior commissure line. The sequence parameters were as follows: time of repetition/time of echo 17,000/95.7 ms; slices 66, without gaping; voxel size 1 $\times$ 1 $\times$ 2 mm^3; matrix 256 $\times$ 256, field of view = 256 mm; number of excitations = 1).

DTI analysis was performed using the Oxford University FMRIB Software Library (Jenkinson et al., 2012). As shown in our previous report (Tamura et al., 2021b), we first corrected the eddy currents and movements in the DTI data of the patients using "eddy_correct." We then created a brain mask by running "bet" on non-diffusion weighting images and fit the diffusion tensor model using "DTI fit."

Two DTI indices, fractional anisotropy (FA) and mean diffusivity (MD), were calculated. FA has a range of 0–1, and low values indicate poor white matter tract integrity. In contrast, high MD values indicate poor integrity. Subsequently, we applied non-linear registration of FA or MD images onto the probabilistic brain atlas MNI152 to normalize them and extracted the mean FA skeletons of all patients, which represented and ran in the center of each tract. Finally, we projected each patient's FA or MD values onto each skeleton of the tract.

In our previous study (Tamura et al., 2021b), we investigated the association between DTI parameters and sarcopenia using the following seven tracts as regions of interest (ROI): lATR, right ATR (rATR), forceps minor of the corpus callosum (FM), left IFOF (lIFOF), rIFOF, and left and right superior longitudinal fasciculus (SLF). This was because we hypothesized that executive and visuospatial dysfunction might be involved in the progression of sarcopenia and frailty, and it had been shown that these tracts might have some roles in these functions. Therefore, we selected these seven tracts as ROIs in this study.

Statistical analyses

First, comparisons of DTI abnormalities between those who became frail during the follow-up period and those who did not were performed using the Mann–Whitney U test. Next, Cox proportional hazard models were used with the incidence of frailty according to each diagnostic criterion (mCHS and KCL) set as the objective variable and the following explanatory variables: FA or MD values plus age, sex, and body mass index (BMI) (Model 1); Model 1 plus hemoglobin A1c level (HbA1c), systolic blood pressure (sBP), and physical activity (Model 2); and Model 2 plus MMSE score (Model 3); BMI, physical activity and MMSE were significantly different in the Mann–Whitney U test in at least one of the frailty criteria. sBP is associated with deterioration of DTI markers in some tracts (Rosano et al., 2015; Tamura et al., 2021a). High HbA1c is shown to be associated with the incidence of frailty. Finally, we performed a stratified analysis of Model 3 based on the diagnostic status

of DM. Estimated frailty-free survival curves were drawn by the Kaplan–Meier method and differences in survival between high MD and low MD groups in the lATR were evaluated with the log-rank test.

Statistical significance was set at $P < 0.05$. All analyses were performed using SPSS Version 20.0 (IBM Corp., Armonk, NY, USA).

Results

Baseline characteristics of the patients

The baseline characteristics of the selected patients, those who were not frail according to either the mCHS or KCL criteria at baseline and those with follow-up data of the frailty status, are shown in Table 1. The prevalence of diabetes, hypertension, and dyslipidemia at baseline were 62.5%, 73.9%, and 73.4%, respectively.

Comparisons of DTI abnormalities at baseline between those who did and did not develop frailty

Among the 137 patients who were not diagnosed as frail at baseline according to the mCHS criteria, 46 (34%) developed frailty during a median follow-up period of 3.0 years. Similarly, among 154 patients who were not diagnosed as frail according to the KCL criteria, 48 (31%) developed frailty during the median follow-up period of 3.0 years. Among the 107 patients who were not diagnosed as frail according to both criteria and whose subsequent frailty status according to both criteria were available, 12 (11%) developed frailty according to the mCHS criteria alone, 10 (9%) developed frailty according to the KCL criteria alone, and 19 (18%) developed frailty according to both criteria.

The results of the Mann–Whitney U test are shown in Table 2. Among the patients without frailty at baseline according to mCHS criteria, those who developed frailty showed lower baseline FA values in many tracts, including the FA in the rATR, FM, lIFOF, and bilateral SLF, and higher MD values in all seven tested tracts, compared with patients who did not develop frailty. Similarly, those who developed frailty according to the KCL criteria showed significantly lower baseline FA and higher MD values in the bilateral ATR than those who did not develop frailty. In both diagnostic criteria, the tracts with the most significant differences were the MD values in the rATR, with values of 8.06×10^{-4} vs. 7.80×10^{-4}, $p = 0.001$ and 8.06×10^{-4} vs. 7.81×10^{-4}, $p = 0.001$ (mCHS and KCL, respectively).

Cox proportional hazard models for the incidence of frailty

The results are shown in Table 3A (mCHS-defined frailty) and Table 3B (KCL-defined frailty). The MD value in the lATR

TABLE 1 Baseline clinical characteristics of the patients.

	Total (n = 184)	Diabetes Mellitus (-) (n = 69)	Diabetes Mellitus (+) (n = 115)	p
Age	77 [74–81]	76 [74–82]	77 [75–81]	0.437
Women (%)	64.1	71.0	60.0	0.132
BMI (kg/m^2)	23.2 [21.1–25.7]	22.5 [20.5–25.5]	23.2 [21.1–25.7]	0.107
Hypertension (%)	73.9	75.4	73.0	0.729
Diabetes Mellitus (%)	62.5	0	100	–
Dyslipidemia (%)	73.4	66.7	77.4	0.111
Cardiovascular disease (%)	15.3	10.3	18.3	0.148
sBP (mmHg)	130 [119–140]	129 [116–139]	130 [121–140]	0.408
dBP (mmHg)	74 [67–82]	73 [67–82]	74 [68–82]	0.859
Alb (g/dL)	4.0 [3.9–4.2]	4.0 [3.8–4.2]	4.0 [3.9–4.2]	0.496
Hb (g/dL)	13.0 [12.4–13.8]	13.1 [12.5–14.0]	13.0 [12.3–13.8]	0.571
HbA1c (%)	6.5 [5.9–7.2]	5.9 [5.7–6.1]	7.0 [6.6–7.5]	**<0.001**
GA/HbA1c	2.73 [2.58–2.88]	2.69 [2.58–2.82]	2.81 [2.56–2.97]	0.051
TG (mg/dL)	115 [82–153]	112 [79–170]	116 [83–147]	0.634
LDL-C (mg/dL)	107 [89–125]	117 [96–139]	102 [87–120]	**<0.001**
HDL-C (mg/dL)	57 [48–69]	60 [51–71]	55 [46–66]	0.058
Cre (mg/dL)	0.80 [0.67–1.03]	0.77 [0.66–0.95]	0.81 [0.70–1.05]	0.174
eGFR (mL/min/1.73 m^2)	58.2 [47.9–69.5]	60.2 [48.6–69.9]	56.6 [47.6–69.2]	0.434
MMSE	29 [27–29]	29 [28–29]	28 [26–29]	0.074
Suspected MCI (%)	75.5	71.0	78.3	0.268
EE (metabolic equivalents × min/week)	1386 [594–2772]	1386 [811–2354]	1386 [491–2777]	0.751
mCHS score	1 [0–2]	1 [0–2]	1 [0–2]	0.346
KCL score	5 [3–7]	5 [3–7]	5 [3–7]	0.964

Data are presented as median [25–75 percentile]. BMI, body mass index; sBP, systolic blood pressure; dBP, diastolic blood pressure; Alb, albumin level; Hb, hemoglobin level; GA, Glucoalbumin level; HbA1c, hemoglobin A1c level; TG, triglyceride level; LDL-C, low-density lipoprotein cholesterol level; HDL-C, high-density lipoprotein cholesterol level; Cre, creatinine level; eGFR, estimated glomerular filtration rate; MMSE, Mini-Mental State Examination score; MCI, mild cognitive impairment; EE, energy expenditure; mCHS, modified Cardiovascular Health Study criteria; KCL, Kihon Checklist. MCI was defined as MMSE ≥ 24 and MoCA-J ≤ 25. Bold values indicate statistical significance $p < 0.05$.

10.3389/fnagi.2022.912972

TABLE 2 Differences in FA and MD values in the tested tracts at baseline between the patients with and without frailty as defined using the mCHS and KCL.

Tracts		mCHS (n = 137)		KCL (n = 154)	
		Frailty (+) (n = 46)	**Frailty (-) (n = 91)**	**Frailty (+) (n = 48)**	**Frailty (-) (n = 106)**
lATR	FA	0.529 (0.501–0.546)	0.533 (0.516–0.557)	**0.524 (0.503–0.542)**	**0.538 (0.518–0.559)****
	MD($\times 10^{-4}$)	**7.83 (7.43–8.16)**	**7.45 (7.16–7.87)****	**7.69 (7.29–8.21)**	**7.44 (7.18–7.77)****
rATR	FA	**0.501 (0.479–0.520)**	**0.510 (0.492–0.529)***	**0.502 (0.487–0.521)**	**0.511 (0.496–0.530)***
	MD($\times 10^{-4}$)	**8.06 (7.93–8.66)**	**7.80 (7.51–8.19)****	**8.06 (7.76–8.39)**	**7.81 (7.55–8.09)****
FM	FA	**0.551 (0.536–0.568)**	**0.561 (0.545–0.581)***	0.561 (0.538–0.577)	0.564 (0.548–0.584)
	MD($\times 10^{-4}$)	**8.14 (7.89–8.36)**	**7.94 (7.68–8.20)****	7.92 (7.65–8.28)	7.96 (7.73–8.21)
lIFOF	FA	**0.527 (0.506–0.547)**	**0.539 (0.516–0.562)***	0.538 (0.507–0.556)	0.540 (0.520–0.563)
	MD($\times 10^{-4}$)	**8.24 (7.94–8.47)**	**8.04 (7.77–8.29)***	8.04 (7.84–8.39)	8.02 (7.80–8.33)
rIFOF	FA	0.524 (0.508–0.538)	0.535 (0.511–0.554)	0.525 (0.508–0.556)	0.540 (0.513–0.553)
	MD($\times 10^{-4}$)	**8.22 (8.02–8.42)**	**7.99 (7.76–8.35)****	8.10 (7.81–8.35)	8.02 (7.77–8.37)
lSLF	FA	**0.506 (0.482–0.528)**	**0.521 (0.503–0.537)***	0.520 (0.501–0.540)	0.522 (0.505–0.537)
	MD($\times 10^{-4}$)	**7.63 (7.40–7.93)**	**7.52 (7.27–7.71)***	7.49 (7.27–7.72)	7.45 (7.25–0.770)
rSLF	FA	**0.511 (0.481–0.533)**	**0.527 (0.501–0.540)***	0.523 (0.505–0.541)	0.527 (0.501–0.540)
	MD($\times 10^{-4}$)	**7.73 (7.36–8.00)**	**7.52 (7.27–7.69)****	7.55 (7.28–7.77)	7.51 (7.24–0.770)

FA, fractional anisotropy; MD, mean diffusivity; mCHS, modified Cardiovascular Health Study criteria; KCL, Kihon Checklist; lATR, left anterior thalamic radiation; rATR, right anterior thalamic radiation; FM, forceps minor, lIFOF, left inferior fronto-occipital fasciculus; rIFOF, right inferior fronto-occipital fasciculus; lSLF, left superior longitudinal fasciculus; rSLF, right superior longitudinal fasciculus; Data are presented as median (25–75 percentile). *$p < 0.05$, **$p < 0.01$ vs. Frailty(+). Bold values indicate statistical significance $p < 0.05$.

TABLE 3A Cox proportional hazard models of the DTI values for the incidence of frailty according to the mCHS (n = 137).

Tract		Model 1		Model 2		Model 3	
		HR (95%CI)	**p**	**HR (95%CI)**	**p**	**HR (95%CI)**	**p**
lATR	FA	1.481 (0.514–4.268)	0.467	1.522 (0.505–4.583)	0.456	1.460 (0.492–4.331)	0.496
	MD	**2.137 (1.247–3.661)**	**0.006**	**2.161 (1.201–3.887)**	**0.010**	**2.029 (1.134–3.631)**	**0.017**
rATR	FA	1.226 (0.385–3.902)	0.730	1.298 (0.412–4.091)	0.656	1.365 (0.426–4.373)	0.600
	MD	1.462 (0.946–2.260)	0.087	1.551 (0.954–2.522)	0.077	1.615 (0.985–2.649)	0.058
FM	FA	1.173 (0.359–3.930)	0.791	1.012 (0.314–3.262)	0.984	1.345 (0.420–4.308)	0.618
	MD	1.614 (0.741–3.515)	0.229	1.395 (0.626–3.109)	0.415	1.611 (0.753–3.447)	0.219
lIFOF	FA	1.252 (0.456–3.438)	0.662	1.463 (0.511–4.188)	0.479	1.614 (0.565–4.611)	0.372
	MD	1.040 (0.495–2.184)	0.918	1.163 (0.535–2.528)	0.704	1.295 (0.595–2.819)	0.514
rIFOF	FA	0.861 (0.364–2.037)	0.733	1.087 (0.445–2.657)	0.855	1.206 (0.475–3.061)	0.694
	MD	0.982 (0.733–1.316)	0.904	1.066 (0.790–1.438)	0.677	1.122 (0.826–1.524)	0.462
lSLF	FA	1.037 (0.451–2.383)	0.932	1.180 (0.526–2.645)	0.688	1.358 (0.590–3.129)	0.472
	MD	0.879 (0.503–1.536)	0.650	0.892 (0.497–1.602)	0.703	0.994 (0.552–1.792)	0.985
rSLF	FA	1.063 (0.544–2.078)	0.859	1.161 (0.584–2.308)	0.671	1.262 (0.625–2.548)	0.516
	MD	1.019 (0.568–1.826)	0.950	1.067 (0.587–1.939)	0.832	1.170 (0.637–2.151)	0.612

Model 1: Adjusted for age, sex, and body mass index at baseline. Model 2: Model 1 plus systolic blood pressure, hemoglobin A1c level, and physical activity at baseline. Model 3: Model 2 plus MMSE score at baseline. DTI, diffusion tensor imaging; mCHS, modified Cardiovascular Health Study criteria; FA, fractional anisotropy; MD, mean diffusivity; lATR, left anterior thalamic radiation; rATR, right anterior thalamic radiation; FM, forceps minor, lIFOF, left inferior fronto-occipital fasciculus; rIFOF, right inferior fronto-occipital fasciculus; lSLF, left superior longitudinal fasciculus; rSLF, right superior longitudinal fasciculus; MMSE, Mini-Mental State Examination score; HR, hazard ratio; CI, confidence interval. Hazard ratios are shown per 0.1 decrease for FA and per 1×10^{-4} increase for MD. Bold values indicate statistical significance $p < 0.05$.

showed a significant association with the incidence of frailty according to the mCHS criteria in Models 1, 2, and 3. Similarly, the MD values in the bilateral ATR were significantly associated with the incidence of frailty according to the KCL criteria.

Additionally, the DTI values of the following tracts were associated with the incidence of KCL-defined frailty: the FA in bilateral ATR and rIFOF showed statistically significant associations, while that in the FM and lIFOF showed weak associations. For MD, bilateral SLF also showed association trends. Specifically, the hazard ratios of KCL-defined frailty incidence were 4.26 when the FA of the lATR dropped by 0.1 and 1.75 when the MD of the tract rose by 1×10^{-4}. However, for mCHS-defined frailty, no other tested tracts showed a significant association. In the stratified analysis (Model 3) based on the diagnostic status of DM, results similar to those of the total sample were observed among patients with DM. Contrarily, among those without DM, the association between DTI changes and incidence of frailty was found in none of the tested tracts and with none of the frailty definitions (Table 4).

Finally, we performed Cox regression analysis by adding the history of symptomatic stroke as a confounding factor.

TABLE 3B Cox proportional hazard models of the DTI values for the incidence of frailty according to the KCL criteria (n = 154).

		Model 1		Model 2		Model 3	
Tract		**HR (95%CI)**	**p**	**HR (95%CI)**	**p**	**HR(95%CI)**	**p**
lATR	FA	**4.834 (1.862–12.55)**	**0.001**	**4.576 (1.630–12.85)**	**0.004**	**4.259 (1.561–11.62)**	**0.005**
	MD	**1.918 (1.250–2.942)**	**0.003**	**1.770 (1.143–2.740)**	**0.010**	**1.748 (1.111–2.752)**	**0.016**
rATR	FA	**5.301 (1.486–18.91)**	**0.010**	**3.921 (1.125–13.66)**	**0.032**	**3.959 (1.114–14.07)**	**0.033**
	MD	**2.138 (1.390–3.289)**	**0.001**	**1.893 (1.239–2.893)**	**0.003**	**1.938 (1.252–2.999)**	**0.003**
FM	FA	2.899 (0.929–8.980)	0.067	2.637 (0.780–8.912)	0.119	**3.589 (1.047–12.30)**	**0.042**
	MD	1.267 (0.574–2.800)	0.558	1.175 (0.510–2.711)	0.705	1.536 (0.655–3.601)	0.324
lIFOF	FA	2.489 (0.967–6.404)	0.059	2.485 (0.909–6.799)	0.076	**2.844 (1.039–7.782)**	**0.042**
	MD	1.214 (0.563–2.618)	0.620	1.306 (0.615–2.774)	0.488	1.476 (0.677–3.216)	0.328
rIFOF	FA	**2.927 (1.118–7.664)**	**0.029**	**2.919 (1.086–7.847)**	**0.034**	**3.845 (1.355–10.91)**	**0.011**
	MD	**1.489 (1.018–2.178)**	**0.040**	1.433 (0.996–2.063)	0.053	1.536 (0.655–3.601)	0.324
lSLF	FA	1.591 (0.613–4.133)	0.340	1.753 (0.658–4.672)	0.262	2.038 (0.739–5.621)	0.169
	MD	1.771 (0.930–3.371)	0.082	**1.929 (1.029–3.617)**	**0.040**	**2.350 (1.229–4.494)**	**0.010**
rSLF	FA	1.362 (0.599–3.100)	0.461	1.556 (0.668–3.622)	0.305	1.675 (0.714–3.926)	0.236
	MD	1.450 (0.801–2.627)	0.220	1.799 (0.963–3.360)	0.066	**2.034 (1.075–3.847)**	**0.029**

Model 1: Adjusted for age, sex, and body mass index at baseline. Model 2: Model 1 plus systolic blood pressure, hemoglobin A1c level, and physical activity at baseline. Model 3: Model 2 plus MMSE score at baseline. DTI, diffusion tensor imaging; KCL, Kihon Checklist; FA, fractional anisotropy; MD, mean diffusivity; lATR, left anterior thalamic radiation; rATR, right anterior thalamic radiation; FM, forceps minor, lIFOF, left inferior fronto-occipital fasciculus; rIFOF, right inferior fronto-occipital fasciculus; lSLF, left superior longitudinal fasciculus; rSLF, right superior longitudinal fasciculus; MMSE, Mini-Mental State Examination score; HR, hazard ratio; CI, confidence interval. Hazard ratios are shown per 0.1 decrease for FA and per 1×10^{-4} increase for MD. Bold values indicate statistical significance $p < 0.05$.

The effects of changes in DTI values in some tracts (rATR, FM, and lIFOF) on the incidence of KCL-defined frailty were slightly attenuated in all patients and those with DM. However, the associations of lATR values on the incidence of both mCHS-defined and KCL-defined frailty remained significant even after adding symptomatic stroke as an explanatory variable. **Figure 1** shows the Kaplan–Meier curve of the frailty-free survival rate in patients with high MD ($\geq 7.8 \times 10^{-4}$) and low MD ($<7.8 \times 10^{-4}$) in the lATR. In both frailty criteria, the high MD group showed a significantly higher rate of frailty incidence than the low MD group ($p < 0.01$ in both; log-rank test).

Discussion

In this study, we prospectively investigated the association between white matter changes evaluated using DTI and the incidence of frailty; abnormalities in some tracts, especially the ATR, were found to be involved in the association.

The strength of this study is that, unlike previous studies that included healthy community-dwelling older adults, we used the data of a substantial number of outpatients (>180), most of whom had cardiometabolic diseases, accumulation of which is a strong risk for frailty as well as for white matter deterioration. In our recent cross-sectional study, some early markers for atherosclerosis were associated with changes in DTI indices (Tamura et al., 2021a). It is important to understand how changes in white matter integrity affect the frailty status of these high-risk patients. Furthermore, we evaluated frailty status using the mCHS and KCL. Recently, it has been emphasized that frailty should not just be evaluated using a physical frailty scale, but scales should encompass broad aspects such as cognitive function and social activity since these factors correlate with each other (Ma et al., 2018). The KCL incorporates various domains such as cognitive dysfunction, malnutrition, social isolation, and depression and is associated not only with the number of CHS frailty phenotypes but also with nutritional state, cognitive function, and depressive mood (Satake et al., 2016), which could be a reason for its strong association with DTI changes.

Herein, we found that alterations in DTI values in the ATR at baseline had the strongest influence on the incidence of frailty. Only a few studies have investigated the association between frailty and abnormalities in white matter integrity assessed using DTI. A cross-sectional study (Avila-Funes et al., 2017) and a longitudinal study (Maltais et al., 2020) investigated the associations between lesion localization and outcome; both studies demonstrated that lesions of the anterior limb of the internal capsule were among those responsible for the presence or worsening of frailty. The method of classifying white matter tracts in our study differed from those of the studies mentioned above as we did not directly verify the anterior limb of the internal capsule. Nevertheless, as ATR fibers run through the anterior limb of the internal capsule, our results are partially comparable with their results. We found that changes in MD in the lATR influenced both mCHS- and KCL-defined frailty. The ATR is involved in gait speed and executive function (Poole et al., 2018), indicating that this region is crucial for maintaining functional and cognitive ability. However, since the association between MD values and the incidence of frailty was observed even after adjusting for physical activity and cognitive function, DTI changes could induce frailty *via* an unknown pathway independent of these dysfunctions.

Conversely, we found an association between FA change and outcomes only in KCL-defined frailty and not in mCHS-defined frailty, indicating that changes in MD values in the ATR were

TABLE 4 Cox regression analysis of the DTI values for the incidence of frailty stratified according to diabetes mellitus status.

		mCHS (n = 137)				KCL (n = 154)			
		Diabetes Mellitus (-) (n = 48)		Diabetes Mellitus (+) (n = 89)		Diabetes Mellitus (-) (n = 58)		Diabetes Mellitus (+) (n = 96)	
Tract		**HR (95%CI)**	**p**	**HR (95%CI)**	**p**	**HR (95%CI)**	**p**	**HR (95%CI)**	**p**
lATR	FA	0.159 (0.008–3.321)	0.231	2.084 (0.622–6.989)	0.234	13.71 (0.815–230.5)	0.069	**4.208 (1.217–14.55)**	**0.023**
	MD	1.468 (0.329–6.554)	0.615	**2.231 (1.132–4.396)**	**0.020**	2.684 (0.630–11.43)	0.182	**1.828 (1.098–3.043)**	**0.020**
rATR	FA	0.266 (0.007–10.55)	0.481	1.612 (0.376–6.908)	0.520	3.142 (0.158–62.40)	0.453	**5.830 (1.239–27.44)**	**0.026**
	MD	2.270 (0.481–10.71)	0.300	1.622 (0.907–2.902)	0.103	1.926 (0.485–7.650)	0.352	**2.331 (1.421–3.824)**	**0.001**
FM	FA	0.079 (0.003–2.194)	0.134	3.085 (0.592–16.08)	0.181	10.92 (0.802–148.7)	0.073	3.538 (0.792–15.81)	0.098
	MD	0.343 (0.030–3.946)	0.391	2.897 (0.960–8.738)	0.059	5.457 (0.960–31.03)	0.056	1.190 (0.434–3.265)	0.736
lIFOF	FA	0.394 (0.029–5.315)	0.483	3.200 (0.830–12.34)	0.091	1.717(0.157–18.83)	0.658	**3.926 (1.215–12.69)**	**0.022**
	MD	0.078 (0.003–2.033)	0.125	2.334 (0.843–6.460)	0.103	2.132 (0.359–12.66)	0.405	1.522 (0.601–3.849)	0.375
rIFOF	FA	0.362 (0.022–5.995)	0.478	1.559 (0.505–4.808)	0.440	3.188 (0.296–34.33)	0.339	**5.033 (1.439–17.61)**	**0.011**
	MD	0.076 (0.005–1.183)	0.066	1.163 (0.826–1.637)	0.387	1.917 (0.320–11.48)	0.476	**1.713 (1.153–2.544)**	**0.008**
lSLF	FA	0.386 (0.052–2.872)	0.353	2.547 (0.742–8.740)	0.137	1.650 (0.294–9.268)	0.570	2.720 (0.669–11.06)	0.162
	MD	0.372 (0.064–2.159)	0.270	1.192 (0.566–2.513)	0.644	2.077 (0.645–6.682)	0.220	**2.927 (1.283–6.675)**	**0.011**
rSLF	FA	0.527 (0.087–3.209)	0.487	1.802 (0.701–4.633)	0.221	1.674 (0.380–7.374)	0.496	2.010 (0.626–6.453)	0.241
	MD	0.696 (0.180–2.695)	0.600	1.603 (0.667–3.852)	0.291	1.583 (0.570–4.287)	0.386	**3.325 (1.229–8.993)**	**0.018**

Adjusted for age, sex, body mass index, systolic blood pressure, hemoglobin A1c level, physical activity, and MMSE score at baseline (Model 3). DTI, diffusion tensor imaging; FA, fractional anisotropy; MD, mean diffusivity; mCHS, modified Cardiovascular Health Study criteria; KCL, Kihon Checklist; lATR, left anterior thalamic radiation; rATR, right anterior thalamic radiation; FM, forceps minor, lIFOF, left inferior fronto-occipital fasciculus; rIFOF, right inferior fronto-occipital fasciculus; lSLF, left superior longitudinal fasciculus; rSLF, right superior longitudinal fasciculus; MMSE, Mini-Mental State Examination score; HR, hazard ratio; CI, confidence interval. Hazard ratios are shown per 0.1 decrease for FA and per 1×10^{-4} increase for MD. Bold values indicate statistical significance $p < 0.05$.

10.3389/fnagi.2022.912972

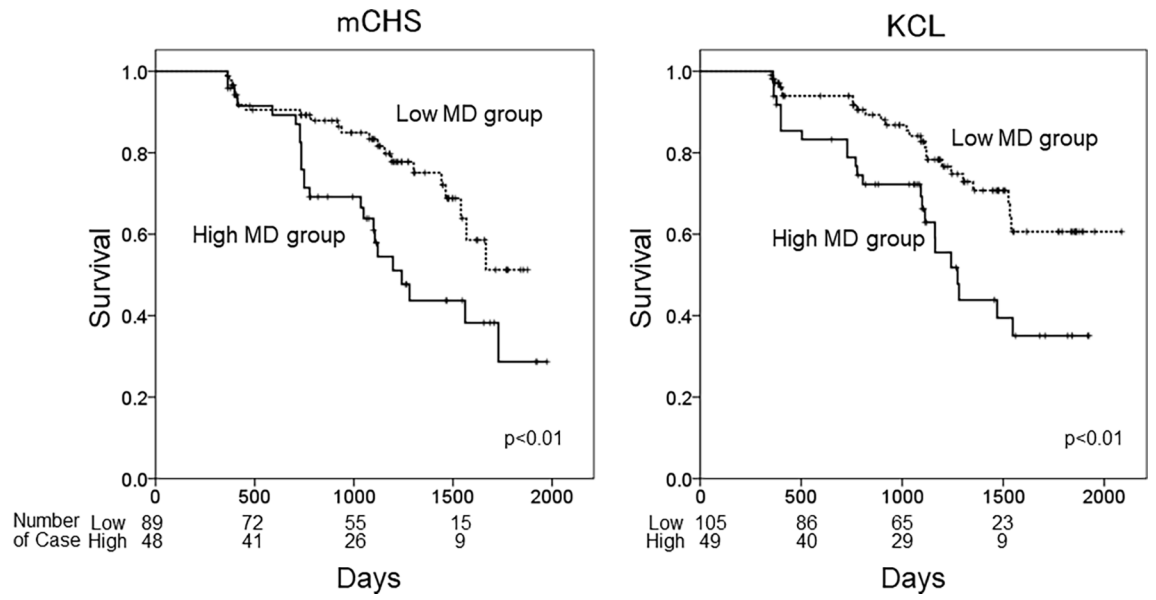


FIGURE 1
Kaplan–Meier curves of frailty-free survival. The patients were categorized into two groups based on frailty defined using the mCHS or KCL and based on the cutoff MD values of the left ATR. High MD group: MD $\geq 7.8 \times 10^{-4}$, low MD group: MD $< 7.8 \times 10^{-4}$. mCHS, modified Cardiovascular Health Study criteria; KCL, Kihon Checklist; lATR, left anterior thalamic radiation; MD, mean diffusivity; rATR, right anterior thalamic radiation.

more strongly associated with CHS-defined frailty than with changes in FA. This is similar to the results obtained by Maltais et al. (2020) probably because the FA values in the crossing-fiber regions are underestimated (Jbabdi et al., 2010). However, it is unclear why the difference was significant only in KCL- and not in mCHS-defined frailty.

In other tracts, the results of our study were mostly compatible with those of Maltais et al. (2020). Specifically, both studies showed a significant association between the alteration of MD in the left SLF and worsening of frailty, although the definition of frailty used was different from that in the mCHS and KCL. As multiple reports have shown that the integrity of the SLF is associated with gait speed (Rosario et al., 2016; Poole et al., 2018), it is natural that abnormalities in this region can lead to frailty. However, we observed only a weak effect of the DTI values of FM on the incidence of frailty, similar to Maltais et al. (2020), who showed no association between FA or MD values of the body or genu of the corpus callosum (CC) and a worsening frailty phenotype. The CC connects the bilateral hemispheres and plays multiple roles. The FM is a group of tracts that extend from the anterior CC (genu) to the cortex of the frontal lobes (Mamiya et al., 2018). It has been indicated that DTI alterations in the FM and the ATR are associated with impaired executive function in young bilingual adults (Mamiya et al., 2018) and that alterations in the genu of the CC are associated with impaired executive function (Zheng et al., 2014) and poor gait performance (de Laat et al., 2011); however, the total effect of lesions in FM on the progression of frailty might be somewhat smaller than those in the ATR or SLF.

We found that low FA and high MD values in the rIFOF were associated with a high incidence of KCL-defined frailty, although the difference was not statistically significant (**Table 2**). The IFOF connects the frontal and occipital lobes (Jou et al., 2011) and is involved in visuospatial cognition (Voineskos et al., 2012), together with superior FOF; Maltais et al. (2020) showed that the MD values of this tract were associated with frailty incidence. Our previous cross-sectional study demonstrated that the abnormalities in the lATR and rIFOF are associated with the prevalence of sarcopenia, especially in patients with DM (Tamura et al., 2021b). These results suggest that alterations in the integrity of the IFOF might induce frailty *via* the impairment of muscle function. Further studies are needed to clarify how changes in the integrity of specific tracts can lead to frailty.

In the stratified analysis, we found associations between the changes in DTI values and the incidence of frailty, specifically in patients with DM and not in those without DM. This could be because patients with DM may have a concurrent etiological background for both white matter changes and the incidence of frailty. Insulin resistance and oxidative stress could be candidate factors; however, we could not evaluate these factors in this study. Individuals with high insulin resistance reportedly show alterations in DTI values in broad white matter tracts (Ryu et al., 2014). Contrarily, insulin resistance is strongly related to and forms a vicious cycle of sarcopenia and sarcopenic obesity (Cleasby et al., 2016), which are key factors for frailty progression. Nonetheless, in those without DM, other physical factors besides

white matter changes could play a more important role in developing frailty.

This study had several limitations. First, the follow-up rate was low. We could not evaluate the frailty status of approximately half of the patients at 2 years. The incidence of frailty might have been higher if those who dropped out had been analyzed. Second, although the general rule was to perform the follow-up assessment annually, some patients could not be evaluated every year. This may have resulted in an overestimation of the period until the incidence of frailty. Third, we did not control for multiple comparisons in multivariate analyses and this might lead to overestimations of some of the results. Fourth, we did not evaluate changes in DTI values. The causative effect of the deterioration of white matter integrity over time on the incidence of frailty could not be clarified. Finally, this study was conducted at a single Japanese facility. Multicenter studies are needed to extrapolate our results to a larger population.

In conclusion, changes in various white matter tracts are associated with the incidence of frailty in older adults, especially those with DM. Changes in DTI values in the ATR were most closely related to the incidence of frailty, and it can be used as an early predictive marker of frailty. Future studies are needed to determine whether diet, exercise, and pharmacologic interventions can improve DTI abnormalities or prevent the progression of frailty in patients at high risk for frailty, as revealed by these DTIs.

Data availability statement

The datasets presented in this article are not readily available due to participants' confidentiality. Requests to access the datasets should be directed to Yoshiaki Tamura: tamurayo@tmghig.jp.

Ethics statement

The studies involving human participants were reviewed and approved by the ethics committee of the Tokyo Metropolitan Geriatric Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YT, JI, and AA designed the study, analyzed the data, and wrote the draft of the manuscript. KS contributed to DTI data collection. YM, FY, RK, KO, KT, and YC contributed to clinical data collection, analysis, and interpretation of data. AT contributed to data interpretation and critically reviewed the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by a grant from the Research Funding for Longevity Sciences (28–30) from the National Center for Geriatrics and Gerontology and the JGS Research Grant Award in Geriatrics and Gerontology from the Japan Geriatrics Society. Also, this work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number JP16H06280, Grant-in-Aid for Scientific Research on Innovative Areas—Platforms for Advanced Technologies and Research Resources "Advanced Bioimaging Support".

Acknowledgments

We thank clinical psychologists Mr. Masahiro Cho, Ms. Kanako Ito, Mr. Hirofumi Hirose, Ms. Manami Ogawa, Ms. Tomomi Kobayashi, and Ms. Aiko Cho for interviewing and performing physical function tests; and Ms. Kazuko Minowa and Ms. Minako Yabuki for data management.

Conflict of interest

AA has received speaker honoraria from Merck Sharp & Dohme, Sumitomo Parma Co. Ltd., Tanabe Mitsubishi Pharma Corporation, Ono Pharmaceutical Co. Ltd., Takeda Pharmaceutical Co. Ltd., and Novo Nordisk Pharma Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Assaf, Y., and Pasternak, O. (2008). Diffusion tensor imaging (dti) based white matter mapping in brain research: a review. *J. Mol. Neurosci.* 34, 51–61. doi: 10.1007/s12031-007-0029-0

Avila-Funes, J. A., Pelletier, A., Meillon, C., Catheline, G., Periot, O., Trevin, O., et al. (2017). Vascular cerebral damage in frail older adults: the amimage study. *J. Gerontol. A. Biol. Sci. Med. Sci.* 72, 971–977. doi: 10.1093/gerona/glw347

Cleasby, M. E., Jamieson, P. M., and Atherton, P. J. (2016). Insulin resistance and sarcopenia: mechanistic links between common co-morbidities. *J. Endocrinol.* 229, R67–81. doi: 10.1530/JOE-15-0533

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. Ē., Booth, M. L., Ainsworth, B. E., et al. (2003). International physical activity questionnaire: 12 country reliability and validity. *Med. Sci. Sports Exerc.* 35, 1381–1395. doi: 10.1249/01.MSS.0000078924.61453.FB

de Laat, K. F., Tuladhar, A. M., van Norden, A. G., Norris, D. G., Zwiers, M. P., and de Leeuw, F. E. (2011). Loss of white matter integrity is associated with gait disorders in cerebral small vessel disease. *Brain.* 134, 73–83. doi: 10.1093/brain/awq343

Ensrud, K. E., Ewing, S. K., Taylor, B. C., Fink, H. A., Cawthon, P. M., Stone, K. L., et al. (2008). Comparison of 2 frailty indexes for prediction of falls, disability, fractures and death in older women. *Arch. Intern. Med.* 168, 382–389. doi: 10.1001/archinternmed.2007.113

Fried, L. P., Tangen, C. M., Walston, J., Newman, A. B., Hirsch, C., Gottdiener, J., et al. (2001). Frailty in older adults: evidence for a phenotype. *J. Gerontol A. Biol. Sci. Med. Sci.* 56, M146–M156. doi: 10.1093/gerona/56.3.m146

Gao, K., Li, B. L., Yang, L., Zhou, D., Ding, K. X., Yan, J., et al. (2021). Cardiometabolic diseases, frailty and healthcare utilization and expenditure in community-dwelling Chinese older adults. *Sci. Rep.* 11:7776. doi: 10.1038/s41598-021-8744-z

Hanlon, P., Fauré, I., Corcoran, N., Butterly, E., Lewsey, J., Mcallister, D., et al. (2020). Frailty measurement, prevalence, incidence and clinical implications in people with diabetes: a systematic review and study-level meta-analysis. *Lancet Healthy Longev.* 1, e106–e116. doi: 10.1016/S2666-7568(20)30014-3

Jbabdi, S., Behrens, T. E. J., and Smith, S. M. (2010). Crossing fibers in tract-based spatial statistics. *Neuroimage.* 49, 249–256. doi: 10.1016/j.neuroimage.2009.08.039

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., and Smith, S. M. (2012). Fsl. *NeuroImage.* 62, 782–790. doi: 10.1016/j.neuroimage.2011.09.015

Jou, R. J., Mateljevic, N., Kaiser, M. D., Sugrue, D. R., Volkmar, F. R., and Pelphrey, K. A. (2011). Structural neural phenotype of autism: preliminary evidence from a diffusion tensor imaging study using tract based spatial statistics. *Am. J. Neuroradiol.* 32, 1607–1613. doi: 10.3174/ajnr.A2558

Koyama, S., Otobe, Y., Suzuki, M., Kimura, Y., Ishiyama, D., Kojima, I., et al. (2022). Relationship between the kihon checklist and all-cause hospitalization among community-dwelling older adults. *Geriatr. Gerontol. Int.* 22, 132–137. doi: 10.1111/ggi.14331

Ma, L., Sun, F., and Tang, Z. (2018). Social frailty is associated with physical functioning, cognition and depression and predicts mortality. *J. Nutr. Health Aging* 22, 989–995. doi: 10.1007/s12603-018-1054-0

Maltais, M., de Souto Barreto, P., Perus, L., Mangin, J. F., Grigis, A., Chupin, M., et al. (2020). Prospective associations between diffusion tensor imaging parameters and frailty in older adults. *J. Am. Geriatr. Soc.* 68, 1050–1055. doi: 10.1111/jgs.16343

Mamiya, P. C., Richards, T. L., and Kuhl, P. K. (2018). Right forceps minor and anterior thalamic radiation predict executive function skills in young bilingual adults. *Front. Psychol.* 9:118. doi: 10.3389/fpsyg.2018.00118

Poole, V. N., Wooten, T., Iloputaife, I., Milberg, W., Esterman, M., and Lipsitz, L. A. (2018). Compromised prefronta structure and function are associated with slowe walking in older adults. *Neuroimage Clin.* 20, 620–626. doi: 10.1016/j.nicl.2018.08.017

Rosano, C., Abebe, K. Z., Aizenstein, H. J., Boudreau, R., Jennings, J. R., Venkatraman, V., et al. (2015). Longitudinal systolic blood pressure characteristics and integrity of white matter tracts in a cohort of very old black and white adults. *Am. J. Hypertens.* 28, 326–334. doi: 10.1093/ajh/hpu134

Rosario, B. L., Rosso, A. L., Aizenstein, H. J., Harris, T., Newman, A. B., Satterfield, S., et al. (2016). Cerebral white matter and slowgait: contribution of hyperintensities and normal-appearing parenchyma. *J. Gerontol. A.* 71, 968–973. doi: 10.1093/gerona/glv224

Ryu, S. Y., Coutu, J. P., Rosas, H. D., and Salat, D. H. (2014). Effects of insulin resistance on white matter microstructure in middle-aged and older adults. *Neurology* 82, 1862–1870. doi: 10.1212/WNL.0000000000000452

Satake, S., and Arai, H. (2020). The revised Japanese version of the cardiovascular health study criteria (revised J-CHS criteria). *Geriatr. Gerontol. Int.* 20, 992–993. doi: 10.1111/ggi.14005

Satake, S., Senda, K., Hong, Y. J., Miura, H., Endo, H., Sakurai, T., et al. (2016). Validity of the kihon checklist for assessing frailty status. *Geriatr. Gerontol. Int.* 16, 709–715. doi: 10.1111/ggi.12543

Satake, S., Shimada, H., Yamada, M., Kim, H., Yoshida, H., Gondo, Y., et al. (2017a). Prevalence of frailty among community-dwellers and outpatients in Japan as defined by the Japanese version of the Cardiovascular Health Study criteria. *Geriatr. Gerontol. Int.* 17, 2629–2634. doi: 10.1111/ggi.13129

Satake, S., Shimokata, H., Senda, K., Kondo, I., and Toba, K. (2017b). Validity of total Kihon checklist score for predicting the incidence of 3-year dependency and mortality in a community-dwelling older population. *J. Am. Med. Dir. Assoc.* 18, 552.e1–552.e6. doi: 10.1016/j.jamda.2017.03.013

Tamura, Y., Ishikawa, J., Fujiwara, Y., Tanaka, M., Kanazawa, N., Chiba, Y., et al. (2018). Prevalence of frailty, cognitive impairment and sarcopenia in outpatients with cardiometabolic disease in a frailty clinic. *BMC Geriatr.* 18:264. doi: 10.1186/s12877-018-0955-4

Tamura, Y., Kimbara, Y., Yamaoka, T., Sato, K., Tsuboi, Y., Kodera, R., et al. (2017). White matter hyperintensity in elderly patients with diabetes mellitus is associated with cognitive impairment, functional disability and a high glycoalbumin/glycohemoglobin ratio. *Front. Aging. Neurosci.* 9:220. doi: 10.3389/fnagi.2017.00220

Tamura, Y., Shimoji, K., Ishikawa, J., Matsuo, Y., Watanabe, S., Takahashi, H., et al. (2021a). Subclinical Atherosclerosis vascular risk factors and white matter alterations in diffusion tensor imaging findings of older adults with cardiometabolic diseases. *Front. Aging. Neurosci.* 13:712385. doi: 10.3389/fnagi.2021.712385

Tamura, Y., Shimoji, K., Ishikawa, J., Tachibana, A., Kodera, R., Oba, K., et al. (2021b). Associations between sarcopenia and white matter alterations in older adults with diabetes mellitus: a diffusion tensor imaging study. *J. Diabetes Investig.* 12, 633–640. doi: 10.1111/jdi.13379

Voineskos, A. N., Rajji, T. K., Lobaugh, N. J., Miranda, D., Shenton, M. E., Kennedy, J. L., et al. (2012). Age related decline in white matter tract integrity and cognitive performance: a dti tractography and structural equation modeling study. *Neurobiol. Aging.* 33, 21–34. doi: 10.1016/j.neurobiolaging.2010.02.009

Zheng, Z., Shemmassian, S., Wijekoon, C., Kim, W., Bookheimer, S., and Yand Pouratian, N. (2014). Dti correlates of distinct cognitive impairments in Parkinson's disease. *Hum. Brain Mapp.* 35, 1325–1333. doi: 10.1002/hbm.22256

frontiers | Frontiers in Aging Neuroscience

TYPE Original Research
PUBLISHED 01 September 2022
DOI 10.3389/fnagi.2022.900523

OPEN ACCESS

EDITED BY
Ondrej Bezdicek,
Charles University, Czechia

REVIEWED BY
Chunbo Li,
Shanghai Jiao Tong University, China
Shiwei Duan,
Zhejiang University City College, China
Xiaoniu Liang,
Fudan University, China

*CORRESPONDENCE
Ying Yu
yuying@sumhs.edu.cn
Qi Guo
guoqijp@gmail.com

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Alzheimer's Disease and Related Dementias,
a section of the journal
Frontiers in Aging Neuroscience

RECEIVED 20 March 2022
ACCEPTED 15 August 2022
PUBLISHED 01 September 2022

CITATION
Liu Y, Yu X, Han P, Chen X, Wang F, Lian X, Li J, Li R, Wang B, Xu C, Li J, Zheng Y, Zhang Z, Li M, Yu Y and Guo Q (2022) Gender-specific prevalence and risk factors of mild cognitive impairment among older adults in Chongming, Shanghai, China. *Front. Aging Neurosci.* 14:900523. doi: 10.3389/fnagi.2022.900523

Gender-specific prevalence and risk factors of mild cognitive impairment among older adults in Chongming, Shanghai, China

Yuewen Liu[1,2†], Xing Yu[1†], Peipei Han[1], Xiaoyu Chen[1], Feng Wang[3], Xuan Lian[3], Jiayu Li[3], Ruijin Li[3], Beibei Wang[3], Chunliu Xu[3], Junxue Li[4], Yaqing Zheng[5], Ziwei Zhang[5], Ming Li[2], Ying Yu[1*] and Qi Guo[2*]

[1]Shanghai University of Medicine and Health Sciences, Shanghai, China, [2]Shanghai University of Medicine and Health Sciences Affiliated Zhoupu Hospital, Shanghai, China, [3]Shanghai University of Traditional Chinese Medicine, Shanghai, China, [4]Shanghai Health Rehabilitation Hospital, Shanghai, China, [5]Fujian Provincial Hospital, Fuzhou, China

Objective: This study explores the gender differences in the prevalence of mild cognitive impairment (MCI) and the correlation between multiple influencing factors.

Materials and methods: The sample was comprised of 1325 relatively healthy participants aged ≥ 60 years in a Shanghai community-dwelling (557 males and 768 females). Cognitive function was assessed by Mini-Mental State Examination (MMSE). The Instrumental Activities of Daily Living (IADL) scale was used to assess the activities of daily living.

Results: The overall prevalence of MCI was 15.2%, with 10.2% in men and 18.9% in women. In older male subjects, those with higher the Geriatric Depression Scale (GDS) scores [odds ratio (OR) = 1.07, 95% confidence interval (CI) = 1.01–1.14] and hypertension (OR = 2.33, 95% CI = 1.15–4.73) had a higher risk of MCI. female subjects who were illiterate (OR = 2.95, 95% CI = 1.82–4.78), had a farming background (OR = 1.69, 95% CI = 1.05–2.72), and a history of stroke (OR = 1.96, 95% CI = 1.07–3.59) had a higher risk of MCI, but this was not true for males. However, Male subjects who never smoked were less likely to have MCI (OR = 0.22, 95% CI = 0.09–0.54). Additionally, the prevalence of MCI was lower in older women with high grip strength (OR = 0.96, 95% CI = 0.92–0.99) and hyperlipidemia (OR = 0.45, 95% CI = 0.22–0.96).

Conclusion: The prevalence of MCI was higher in the population of elderly women compared to men. Moreover, it was found that members with MCI tended to having higher GDS scores, smoking, and hypertension; whereas a history of farming, illiteracy, stroke, grip strength, and hyperlipidemia were correlated with MCI in women.

KEYWORDS

mild cognitive impairment, prevalence, risk factors, gender differences, older adults

Introduction

China has 249 million people aged 60 and over, which accounts for 17.9% of the total population (National Bureau of Statistics of China, 2021), indicating a high prevalence of mild cognitive impairment (MCI). With the increase in the aging population worldwide, the number of patients with cognitive impairment is also increasing, putting a heavy burden on families, communities, and health care systems. Accordingly, cognitive impairment has become a global public health problem.

Due to the lack of effective treatment, early intervention is considered the most cost-effective way to manage dementia (Livingston et al., 2017). Since MCI is considered the transition stage between undamaged cognitive function and dementia (Winblad et al., 2004), there has been a consensus to focus the main intervention on this population to prevent dementia. However, a few studies investigating the prevalence of MCI in China have inconsistent results, which are estimated to range from 4.5 to 21.5% (Li et al., 2011; Song et al., 2021). These inconsistencies require further study to arrive at more accurate estimates. Although previous studies have analyzed the association between modifiable risk factors and MCI stratified by gender (Zhang et al., 2019; Fu et al., 2020), this may affect the exposure of MCI risk factors and the prevalence of MCI due to China's vast territory, people's complex lifestyles, extended life span, different diagnostic criteria, and various screening methods. It is estimated that more than 70% of the elderly in China live in suburban communities. Shanghai is one of the most populous megacities in China. Studying the prevalence and influencing factors of MCI in its suburban counties can provide supporting data for the prevention, diagnosis, care, and treatment of MCI. Moreover, the current research conclusions on the gender differences in the prevalence and influencing factors of MCI are still inconsistent. No gender difference in the prevalence of MCI was present in other countries (Ganguli et al., 2013; Au et al., 2017; Overton et al., 2019). The difference is that the data of China show that the prevalence of MCI in women is higher than that in men (Fu et al., 2020). Understanding the gender differences of MCI may further understand the etiology and prevention of dementia. Attention to the prevalence of MCI and effective risk factor prevention strategies will help reduce the incidence rate of MCI and subsequent dementia.

The authors' previous studies have shown that both sarcopenia (Chen et al., 2021) and obesity (Ma et al., 2021) are associated with cognitive impairment. Therefore, this study explores the gender differences in the prevalence of MCI and the correlation between multiple influencing factors. By strengthening the identification of MCI, its related influencing factors can be examined in depth to provide a theoretical basis for the early prevention and identification of MCI to establish an early warning model of MCI.

Materials and methods

Study population

The final analytic sample consisted of 1,325 participants ($\geq$ 60 years old) after excluding 38 individuals (3 had incomplete MMSE data; 4 were unable to perform the physical performance test; 31participants had missing data on covariates or outcomes.). All subjects were invited to participate in a comprehensive geriatric assessment and cognitive function assessment in Chongming District of Shanghai in 2019 and 2020. The exclusion criteria were as follows: (1) severe cognitive impairment, dementia, mental illness or other neurodegenerative diseases; (2) Inability to communicate with researchers or unwillingness to give informed consent; (3) people with deafness or blindness and cannot complete the assessment; (4) those who cannot complete the grip strength, Timed Up and Go Test (TUGT), and 4-meter walking test. All participants provided informed consent prior to participation. If the participant was illiterate, the informed consent of its legal representative would be sought.

Covariates

Data on sociodemographic characteristics, behavioral characteristics, and disease history have previously been described (through face-to-face questions) (Liu et al., 2021). The questionnaire included questions about age, sex, height, body weight, marital status, illiteracy, living habits (alone or with others), sleep duration, smoking habits (current smoker, never smoking, and past smoker), drinking habits (drinking daily, occasional drinking, past drinking, and never drinking), and household income ($<$ 1,000, 1,000–3,000, 3,000–5,000, and $>$ 5,000 RMB), Physical activity was assessed using the short form of the International Physical Activity Questionnaire (IPAQ) (Jiang et al., 2009). The depressive symptoms were evaluated by the Geriatric Depression Scale (GDS). Subjects with a score of $\geq$ 11 were considered to have depressive symptoms (Yesavage et al., 1982). Nutrition was evaluated by the Mini Nutritional Assessment-Short Form (MNA-SF) (Kaiser et al., 2009). Physical performance was assessed by grip strength, a 4-m walking test, and the Timed Up and Go Test (TUGT). The details of the measurement method have been described in the authors' previous research (Liu et al., 2021). Disease history included type 2 diabetes mellitus (T2DM), hypertension, hyperlipidemia, stroke, gout, anemia, pulmonary disease, biliary tract disease,

kidney disease, heart disease, osteoarthritis, cancer, and thyroid disease.

Definition of mild cognitive impairment

This study adopted the MCI diagnostic criteria based on Petersen's definitions (Petersen et al., 1999): (1) memory complaint; (2) normal activities of daily living; (3) normal general cognitive function; (4) abnormal memory for age; (5) not demented.

Assessment of cognitive function

Cognitive assessment was completed using the Mini-Mental State Examination (MMSE) (Folstein et al., 1975) by trained investigators. It includes 30 items, the score ranges from 0 to 30 points, with the higher scores indicating better cognitive performance. The use of MMSE to define MCI is consistent with previous studies (Liu et al., 2021), that is, The cut-off points used for cognitive impairment were as follow: $\leq$ 17 for illiterate people, $\leq$ 20 for people with primary school, and $\leq$ 24 for people with middle school or higher (Zhang et al., 1990).The above cut-off points have been proved to be sensitive and effective in the diagnosis of MCI in Chinese elderly population (Zhang et al., 2006).

Assessment of daily activity ability

The daily activity ability was evaluated using the Instrumental Activities of Daily Living scale (IADL) (Lawton and Brody, 1969). It includes 8 items, the score ranges from 0 to 8 points, with the higher scores indicating better daily activity ability. IADL scores $\geq$ 6 indicates normal daily activity ability (Nagamatsu et al., 2013).

Statistical analysis

The following analyses were performed to investigate the prevalence of MCI and the correlation between multiple influencing factors. Normally distributed data were presented as mean $\pm$ standard deviation, while non-normally distributed data were presented as median, with the 25–75% interquartile range given in parentheses. Categorical variables were expressed as a percentage (%). The differences according to the characteristics of cognitive status were analyzed using the *t*-tests, Chi-square tests, and Mann–Whitney *U* tests. Logistic regression analysis was used to analyze the factors associated with MCI. Based on previous research, age, marital status, illiteracy, living conditions, farming, Smoking, BMI, IPAQ, and comorbidity status (hypertension, hyperlipidemia, stroke, diabetes) were considered factors potentially associated with MCI. As such, these factors were included as independent variables in the models of the present study. All the statistical analyses were performed using SPSS version 21.0, and $P < 0.05$ was considered statistically significant.

Ethics

The study was approved by the Ethics Committee of Shanghai University of Medicine and Health Sciences and the methods were carried out in accordance with the principles of the Declaration of Helsinki.

Results

Characteristics of the participants

A total of 1,363 participants were evaluated. Of these, 38were excluded: 3 had incomplete MMSE data; 4 were unable to perform the physical performance test; 31participants had missing data on covariates or outcomes (**Figure 1**). The final analytic sample consisted of 1,325 participants in the study (mean age, 71.79 $\pm$ 5.77 years; 58.0% women), of whom 202 (15.2%) were diagnosed with MCI (57 males and 145 females). Participants were divided into three age groups, 60–69, 70–79, and $\geq$ 80 years, with 58 (10.5%), 89 (14.2%), and 55 (37.7%) people, respectively. The prevalence increased more steeply with age in females than males (**Figure 2**). **Table 1** presents the characteristics of these participants. Compared to participants with normal cognition, participants with MCI were older, had lower daily activity ability, higher GDS scores, sleep longer, less physical activity, and a higher proportion of illiteracy, widows, living alone, agriculture, low income, and worse physical performance. However, there were some differences when stratified by gender. Males with MCI had a higher proportion of hypertension (77.2% versus 61.8%; $P = 0.022$) and a higher score of GDS (6.54 $\pm$ 4.81 versus 5.50 $\pm$ 4.55; $P = 0.003$), but these results were not statistically significant in females ($P > 0.05$). Females with MCI had longer sleep time, lower daily activity levels, and were more likely to be farmers and illiterate.

Univariate and multivariate analysis of associated factors for mild cognitive impairment

Results from univariate and multivariate logistic regression models for factors related to MCI in 557 males and 768 females

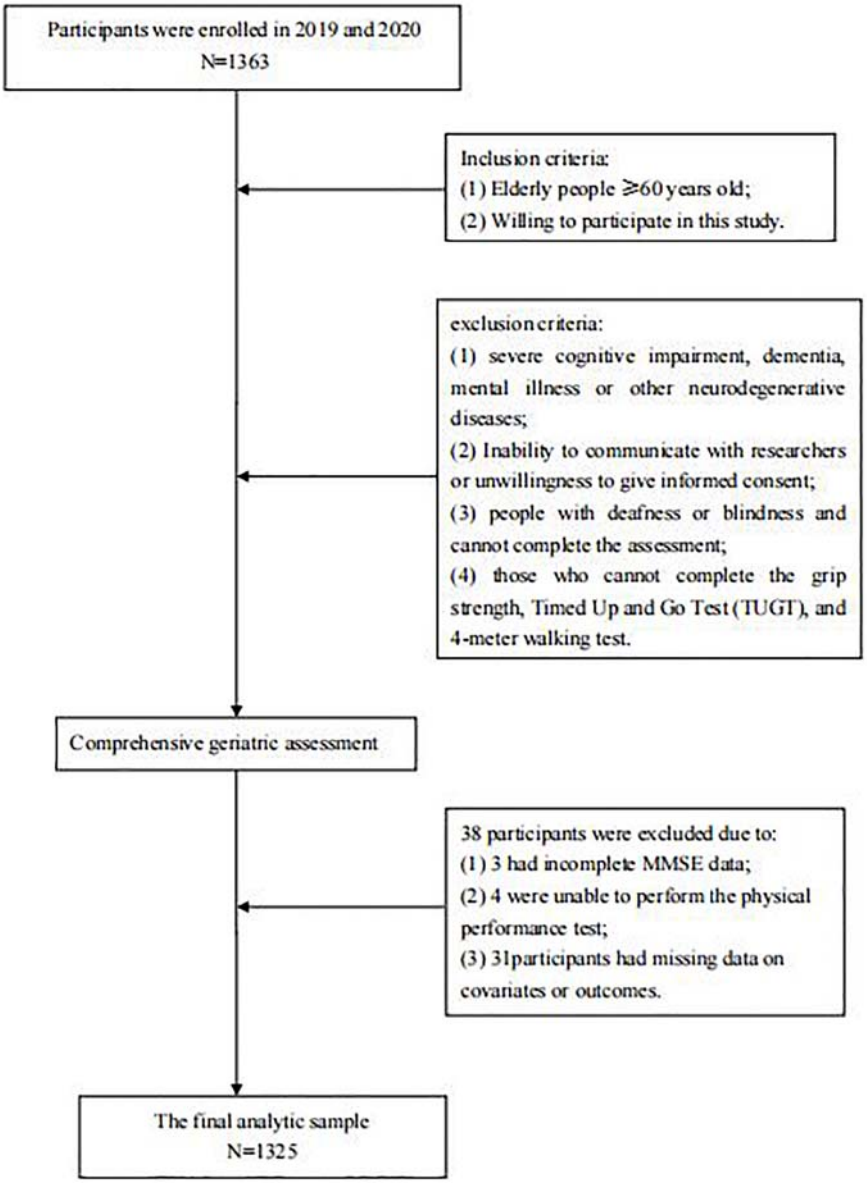


FIGURE 1
Flow chart of the study.

are reported in Tables 2 and 3, respectively. After adjusting for potential confounders, in older male subjects, those with higher GDS scores [odds ratio (OR) = 1.07, 95% confidence interval (CI) = 1.01–1.14] and hypertension (OR = 2.33, 95% CI = 1.15–4.73) were associated with MCI. Male subjects who never smoked were less likely to have MCI than older male subjects who currently smoked (OR = 0.22, 95% CI = 0.09–0.54). Female subjects who were illiterate (OR = 2.95, 95% CI = 1.82–4.78), had a background in farming (OR = 1.69, 95% CI = 1.05–2.72), and had high grip strength (OR = 0.96, 95% CI = 0.92–0.99) were associated with MCI, but this was not true for males. The OR and 95% CI in the adjusted model for the factors statistically significantly associated with MCI were 0.45 (0.22–0.96) and 1.96 (1.07–3.59) for hyperlipidemia and stroke, respectively, in females.

Discussion

This study estimates the prevalence of MCI and the factors associated with MCI in a suburb-dwelling population of elderly persons aged 60 and older in China. The overall prevalence of MCI was 15.2%, with 10.2% in men and 18.9% in women. After adjustment for potential confounders, the GDS scores, smoking, and hypertension were directly associated with the prevalence of MCI in men. Moreover,

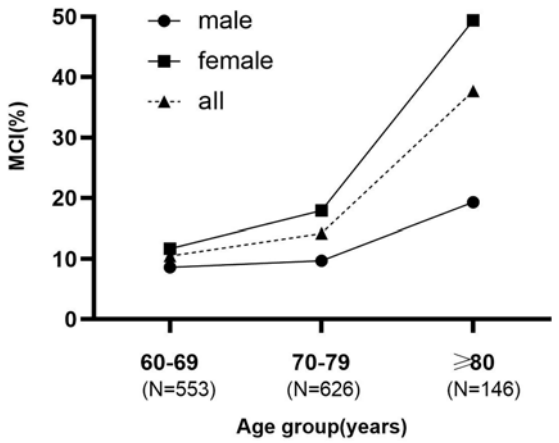


FIGURE 2
The prevalence increased more steeply with age in females than males.

grip strength, being illiterate, having a background in farming, hyperlipidemia, and stroke were associated with the prevalence of MCI in women.

Prevalence of mild cognitive impairment in suburb-dwelling residents

A recent meta-analysis showed that the prevalence of MCI in the Chinese elderly was 15.2% (Deng et al., 2021), which was consistent with the results of the present study. A recent cross-sectional survey on the prevalence of MCI in women showed that the prevalence of MCI was 21.5% (Song et al., 2021), while another study found that the prevalence of cognitive impairment was 26.1% in men and 30.5% in women (Lyu and Kim, 2016). Additionally, a cross-sectional study found that the prevalence of MCI was 19.8% for men and 26.1% for women (Zhang et al., 2019). In the study by Fu et al. (2020), the MCI prevalence was 8.2% in males and 13.1% in females. These large differences in prevalence are likely due to diversity in population structure, screening tools, and diagnostic criteria. These inconsistencies necessitate further study to yield a more accurate estimate.

Modifiable lifestyle and physical performance factors related to mild cognitive impairment in suburb-dwelling residents

The present study found that GDS scores were correlated with the prevalence of MCI. This was consistent with the results of previous studies (Vega and Newhouse, 2014). There is increasing evidence that depression strongly and independently accelerates the progression from MCI to dementia. In individuals with MCI, depressive symptoms mean a 2–4 fold increase in the risk of progression to dementia (Van der Mussele et al., 2014). In men, an association was found between smoking and MCI. This result was consistent with previous work that reported smoking was associated with cognitive impairment (Jia et al., 2020). Regarding the relationship between smoking and MCI, further studies on biomarker assessments of smoking are needed to illustrate the true association between smoking and MCI in the Chinese population. This result also showed that MCI was significantly associated with a background in farming and illiteracy in women. A systematic review showed that those who have a predominantly manual occupation throughout life have a greater risk of cognitive impairment and/or dementia than those in intellectually demanding occupations (Gracia et al., 2016). Illiterate women were more than twice as likely to have MCI than educated women. Previous studies have found that women with little formal education have a higher risk of dementia than men with similar educational backgrounds (Zhang et al., 1990; Launer et al., 1999; Ott et al., 1999). Differences in the quality of education may contribute to an increased risk of dementia in women compared with men. Low-educated women are more likely to have poorer career achievement, lower income, poorer health, less leisure opportunities, and poorer cognitive outcomes than low-educated men (Ott et al., 1999; Sharp and Gatz, 2011). Compared with men, women have less access to education, and when women have access to education, it differs in quality. Zhang et al. (2006) point out that before 1950, only women belonging to the highest social classes could receive education in China. Additionally, a relationship was found between MCI and grip strength in women, as was found in previous studies (Atkinson et al., 2010). This simple measure may be useful for future research on the relationship between cognitive function and physical performance.

Specific diseases associated with mild cognitive impairment in suburb-dwelling residents

As demonstrated in the present study, hypertension was associated with MCI in men, but not in women. This finding was in accordance with previous studies (Wu et al., 2016; Fu et al., 2020). The potential gender differences between hypertension and cognition are not fully understood (Iadecola et al., 2016). It may be that the women in the present study were all menopausal. Previous evidence suggests that older postmenopausal women have a lower risk of cardiometabolic disease because of their favorable hormonal and metabolic profile (Regensteiner et al., 2015), which may mitigate cognitive impairment caused by hypertension. However, the

TABLE 1 Baseline characteristics of study participants with normal cognition vs. mild cognitive impairment.

Variables	Total(n = 1325)			Male(n = 557)			Female(n = 768)		
	Normal cognition (n = 1123)	Mild cognitive impairment (n = 202)	*P-value*	Normal cognition (n = 500)	Mild cognitive impairment (n = 57)	*P-value*	Normal cognition (n = 623)	Mild cognitive impairment (n = 145)	*P-value*
Age (year)	71.27 ± 5.38	74.66 ± 6.96	<0.001	71.70 ± 5.48	73.72 ± 7.02	0.001	70.93 ± 5.28	75.03 ± 6.92	<0.001
BMI(kg/m^2)	23.69 ± 3.45	23.68 ± 3.43	0.963	23.40 ± 3.33	23.39 ± 3.83	0.989	23.92 ± 3.53	23.79 ± 3.27	0.680
IADL (score)	7.79 ± 0.63	7,63 ± 0.67	0.001	7.82 ± 0.47	7.68 ± 0.61	0.039	7.77 ± 0.73	7.61 ± 0.69	0.021
MNA-SF (score)	12.79 ± 1.51	12.68 ± 1.44	0.327	12.82 ± 1.44	12.51 ± 1.66	0.127	12.76 ± 1.56	12.74 ± 1.34	0.886
GDS (score)	5.50 ± 4.55	6.81 ± 5.39	<0.001	4.85 ± 4.01	6.54 ± 4.81	0.003	6.02 ± 4.89	6.92 ± 5.61	0.053
Sleep duration(hour)	8.40 ± 1.43	8.79 ± 1.55	0.001	8.43 ± 1.39	8.61 ± 1.49	0.334	8.38 ± 1.45	8.85 ± 1.57	0.001
IPAQ (Met/wk)	5255(2226,10080)	3864(1647,8201)	0.003	4988(2079, 8966)	3222(1533, 8652)	0.167	5639(2373, 10920)	3990(1908, 7989)	0.003
Illiteracy (%)	98(8.7%)	68(33.7%)	<0.001	21(4.2%)	5(8.8%)	0.121	77(12.4%)	63(43.4%)	<0.001
Widowed(%)	191(17.0%)	75(37.1%)	<0.001	41(8.2%)	11(19.3%)	0.006	150(24.1%)	64(44.1%)	<0.001
Living alone (%)	166(14.8%)	61(30.2%)	<0.001	50(10.0%)	14(24.6%)	0.001	116(18.6%)	47(32.4%)	<0.001
Farming (%)	565(50.3%)	135(66.8%)	<0.001	198(39.6%)	23(40.4%)	0.913	367(58.9%)	112(77.2%)	<0.001
Monthly income (%)			<0.001			0.121			0.004
< 1000	70(6.2%)	29(14.4%)		23(4.6%)	6(10.5%)		47(7.5%)	23(15.9%)	
1000-3000	683(60.8%)	131(64.9%)		286(57.2%)	36(63.2%)		397(63.7%)	95(65.5%)	
3000-5000	160(14.2%)	18(8.9%)		80(16.0%)	7(12.3%)		80(12.8%)	11(7.6%)	
>5000	210(18.7%)	24(11.9%)		111(22.2%)	8(14.0%)		99(15.9%)	16(11.0%)	
Smoking (%)			0.014			0.075			0.495
Current smokers	178(15.9%)	28(13.9%)		174(34.8%)	28(49.1%)		4(0.6%)	0(0.0%)	
Never smokers	757(67.4%)	155(76.7%)		140(28.0%)	10(17.5%)		617(99.0%)	145(100%)	
Ever smokers	188(16.7%)	19(9.4%)		186(37.2%)	19(33.3%)		2(0.3%)	0(0.0%)	
Drinking (%)			0.065			0.168			0.126
Daily drinkers	171(15.2%)	29(14.4%)		150(30.0%)	18(31.6%)		21(3.4%)	11(7.6%)	
Occasional drinkers	152(13.5%)	27(13.4%)		91(18.2%)	16(28.1%)		61(9.8%)	11(7.6%)	
Former drinkers	135(12.0%)	12(5.9%)		111(22.2%)	7(12.3%)		24(3.9%)	5(3.4%)	
Never drinkers	665(59.2%)	134(66.3%)		148(29.6%)	16(28.1%)		517(83.0%)	118(81.4%)	
physical performance									
Grip strength (kg)	23.95 ± 8.91	19.10 ± 8.88	<0.001	30.31 ± 8.22	27.01 ± 9.65	0.005	18.85 ± 5.50	15.99 ± 6.28	<0.001
4-meter walking test (m/s)	1.11 ± 0.23	0.94 ± 0.25	<0.001	1.15 ± 0.25	1.02 ± 0.25	0.001	1.07 ± 0.21	0.92 ± 0.25	<0.001
TUGT (s)	9.63 ± 3.31	11.9 ± 4.76	<0.001	9.56 ± 3.36	11.04 ± 3.70	0.002	9.68 ± 3.28	12.24 ± 5.09	<0.001
Diseases (%)									

(Continued)

10.3389/fnagi.2022.900523

TABLE 1 (Continued)

Variables	Total(n = 1325)			Male(n = 557)			Female(n = 768)		
	Normal cognition (n = 1123)	Mild cognitive impairment (n = 202)	*P-value*	Normal cognition (n = 500)	Mild cognitive impairment (n = 57)	*P-value*	Normal cognition (n = 623)	Mild cognitive impairment (n = 145)	*P-value*
Diabetes (%)	178(15.9%)	32(15.8%)	0.997	78(15.6%)	9(15.8%)	0.970	100(16.1%)	23(15.9%)	0.955
Hypertension (%)	683(60.8%)	139(68.8%)	0.031	309(61.8%)	44(77.2%)	0.022	374(60.0%)	95(65.5%)	0.222
Hyperlipidemia (%)	140(12.5%)	17(8.4%)	0.100	56(11.2%)	7(12.3%)	0.811	84(13.5%)	10(6.9%)	0.029
Stroke(%)	114(10.2%)	29(14.4%)	0.076	51(10.2%)	5(8.8%)	0.734	63(10.1%)	24(16.6%)	0.028
Gout (%)	66(5.9%)	10(5.0%)	0.613	43(8.6%)	5(8.9%)	0.934	23(3.7%)	5(3.4%)	0.888
Anemia (%)	51(4.5%)	13(6.4%)	0.248	18(3.6%)	3(5.3%)	0.532	33(5.3%)	10(6.9%)	0.450
Pulmonary disease (%)	96(8.5%)	15(7.4%)	0.596	58(11.6%)	6(10.5%)	0.810	38(6.1%)	9(6.2%)	0.961
Biliary tract disease (%)	166(14.8%)	32(15.8%)	0.697	49(9.8%)	7(12.3%)	0.555	117(18.8%)	25(17.2%)	0.667
Kidney disease (%)	117(10.4%)	12(5.9%)	0.048	70(14.0%)	5(8.8%)	0.273	47(7.5%)	7(4.8%)	0.249
Heart disease (%)	359(32.0%)	73(36.1%)	0.244	142(28.4%)	19(33.3%)	0.436	217(34.8%)	54(37.2%)	0.584
Thyroid disease (%)	44(3.9%)	10(5.0%)	0.494	11(2.2%)	0(0.0%)	0.258	33(5.3%)	10(6.9%)	0.450
Osteoarthritis(%)	179(15.9%)	39(19.3%)	0.235	68(13.6%)	11(19.3%)	0.243	111(17.8%)	28(19.3%)	0.674
Cancer(%)	44(3.9%)	6(3.0%)	0.515	26(86.7%)	4(7.0%)	0565	18(2.9%)	2(1.4%)	0.304

BMI, Body Mass Index; MNA, Mini Nutritional Assessment; TUGT, Time Up and Go Test; GDS, Geriatric Depression Scale; IADL, Instrumental Activity of Daily Living; IPAQ, International Physical Activity Questionnaires.

10.3389/fnagi.2022.900523

TABLE 2 Unadjusted and Adjusted Model for Factors Related to MCI in the male.

Variable	Univariate Odds Ratio (95% CI)	*P-value*	Adjusted Model Odds Ratio (95% CI)	*P-value*
All sample (n = 557)				
Age (year)				
60-69 (n = 221)	1.0 (referent)		1.0 (referent)	
70-79 (n = 279)	1.14(0.62–2.11)	0.678	1.08(0.54–2.15)	0.823
≥ 80 (n = 57)	2.54(1.13–5.71)	0.024	1.44(0.49–4.27)	0.507
IADL	0.63(0.40–0.99)	0.044	0.83(0.47–1.44)	0.501
GDS	1.09(1.03–1.15)	0.004	1.07(1.01–1.14)	0.033
Illiteracy (n = 26)	2.19(0.79–6.06)	0.130	1.74(0.53–5.69)	0.358
Widowed (n = 52)	2.68(1.29–5.56)	0.008	1.15(0.33–4.01)	0.825
Living alone (n = 64)	2.93(1.50–5.73)	0.002	1.86(0.64–5.36)	0.252
Farming (n = 221)	1.03(0.59–1.80)	0.913	0.70(0.36–1.35)	0.283
Monthly income				
< 1,000 (n = 29)	1.0 (referent)		1.0 (referent)	
1,000–2,999 (n = 322)	0.48(0.18–1.26)	0.138	0.44(0.14–1.36)	0.152
3,000–5,000 (n = 87)	0.34(0.10–1.10)	0.071	0.29(0.07–1.16)	0.080
>5,000 (n = 119)	0.28(0.09–0.87)	0.028	0.28(0.07–1.12)	0.071
Smoking				
Current smokers (n = 202)	1.0 (referent)			
Never smokers (n = 150)	0.44(0.21–0.95)	0.035	0.22(0.09–0.54)	0.001
Ever smokers (n = 205)	0.64(0.34–1.18)	0.150	0.47(0.24–0.95)	0.036
Grip strength	0.95(0.92–0.99)	0.005	0.97(0.93–1.01)	0.105
4-meter walking test	0.16(0.05–0.46)	0.001	0.23(0.04–1.23)	0.085
TUGT	1.10(1.03-1.17)	0.005	0.98(0.88–1.09)	0.670
Hypertension (n = 353)	2.09(1.10–3.99)	0.025	2.33(1.15–4.73)	0.020
Hyperlipidemia (n = 63)	1.11(0.48–2.56)	0.811	1.08(0.42–2.79)	0.879
Stroke (n = 56)	0.85(0.32–2.22)	0.734	0.72(0.25–2.12)	0.554

TUGT, time up and go; OR, Odds Ratio; CI, confidence interval; IPAQ, international physical activity questionnaires; Adjusted model: Adjusted for age, BMI, IADL, GDS, IPAQ, Illiteracy, Widowed, Living alone, Farming, Monthly income, Smoking, Grip strength, 4-meter walking test, TUGT, Hypertension, Hyperlipidemia, Stroke, diabetes.

observational studies conducted for the present study could not provide insight into the underlying biological mechanisms. Accordingly, more research should be conducted to explore possible explanations for the underlying gender differences. In addition, similar to previous studies (Kim and Park, 2017; Fu et al., 2020), the present study found no significant correlation between hyperlipidemia and the prevalence of MCI in the elderly. Interestingly, when stratified by gender, the association between hyperlipidemia and MCI was only significant in women, and women with hyperlipidemia had higher cognitive function than women without hyperlipidemia. In the study conducted by Kim and Park (2017), hyperlipidemia was reported as a protective factor for MCI in women. The relationship between hyperlipidemia and MCI prevalence remains controversial. Studies have shown that a reduced risk of cognitive impairment was associated with higher serum cholesterol levels (Lv et al., 2016). In contrast, earlier research found significantly higher cholesterol levels in older adults with cognitive impairment and suggested that cholesterol-lowering treatments, such as lipid-lowering drugs, may have cardiovascular benefits while preventing cognitive impairment (Vance, 2012). However, the underlying mechanism of the interaction between gender and hyperlipidemia and cognitive impairment remains unclear. Further prospective studies are needed to explore the sex differences with the association of hyperlipidemia with MCI. A recent meta-analysis showed that stroke was a risk factor for dementia, which was consistent with the results of the present study (Deng et al., 2021). Similarly, a study examining gender differences in cognitive outcomes after stroke found that women had significantly worse cognitive outcomes than men (Dong et al., 2020). Since only a few previous works investigated the gender differences between stroke and MCI, further research is needed to explore the potential pathological associations.

Strengths and limitations

This study presents several strengths. First, the study was conducted on a relatively large sample of well-characterized suburban elderly men and women who had lived in a particular geographic area for a long time. Second, the participants were from suburban areas, and their lifestyles were more active, which may have been different from subjects in other

TABLE 3 Unadjusted and adjusted model for factors related to MCI in the female.

Variable	Univariate Odds Ratio (95% CI)	*P-value*	Adjusted Model Odds Ratio (95% CI)	*P-value*
All sample (n = 768)				
Age (year)				
60-69 (n = 332)	1.0 (referent)		1.0 (referent)	
70-79 (n = 347)	1.64(1.06–2.52)	0.026	1.02(0.62–1.67)	0.939
≥ 80 (n = 89)	7.35(4.31–12.52)	<0.001	1.81(0.90–3.63)	0.097
IADL	0.79(0.63–0.98)	0.028	1.26(0.93–1.72)	0.142
GDS	1.03(0.99–1.07)	0.054	0.99(0.95–1.03)	0.608
Sleep duration	1.22(1.08–1.38)	0.001	1.05(0.92–1.20)	0.456
IPAQ	0.96(0.92–0.99)	0.001	0.98(0.96–1.09)	0.131
Illiteracy (n = 140)	5.45(3.63–8.18)	<0.001	2.95(1.82–4.78)	<0.001
Widowed (n = 214)	2.49(1.71–3.63)	<0.001	1.50(0.78–2.88)	0.226
Living alone (n = 163)	2.10(1.40–3.13)	<0.001	0.85(0.41–1.72)	0.642
Farming (n = 479)	2.37(1.56–3.60)	<0.001	1.69(1.05–2.72)	0.031
Monthly income				
< 1,000 (n = 70)	1.0 (referent)		1.0 (referent)	
1,000-2,999 (n = 492)	0.49(0.28–0.85)	0.010	0.67(0.35–1.30)	0.235
3,000-5,000 (n = 91)	0.28(0.13–0.63)	0.002	0.53(0.21–1.34)	0.179
>5,000 (n = 115)	0.33(0.16–0.68)	0.003	0.65(0.28–1.55)	0.330
Grip strength	0.91(0.88–0.94)	<0.001	0.96(0.92–0.99)	0.036
4-meter walking test	0.06(0.03–0.13)	<0.001	0.40(0.11–1.42)	0.156
TUGT	1.16(1.11–1.22)	<0.001	1.05(0.98–1.12)	0.142
Hypertension (n = 469)	1.27(0.87–1.85)	0.223	1.15(0.74–1.81)	0.537
Hyperlipidemia (n = 94)	0.48(0.24–0.94)	0.033	0.45(0.22–0.96)	0.038
Stroke (n = 87)	1.76(1.06–2.94)	0.029	1.96(1.07–3.59)	0.030

TUGT, time up and go; OR, Odds Ratio; CI, confidence interval; IPAQ, international physical activity questionnaires; Adjusted model: Adjusted for age, BMI, IADL, GDS, Sleep duration, IPAQ, Illiteracy, Widowed, Living alone, Farming, Monthly income, Smoking, Grip strength, 4-meter walking test, TUGT, Hypertension, Hyperlipidemia, Stroke, diabetes.

areas. However, this study has some limitations that must be addressed. First, because this is a cross-sectional study, a causal relationship of the identified associated factors could not be established. Second, all participants in this study were relatively healthy. We did not include those who were hospitalized, or those who were bedridden with severe illness and could not be tested on-site, and those who refused to participate in the program. This choice could constitute a selective survival and a healthy selection bias. Third, although there were no patients with depression in our population, we found that men with higher GDS scores had worse cognitive function, which we also adjusted as an adjustment variable in multivariate regression. In the future, we will continue to follow up to observe the impact of depression on cognitive function.

Conclusion

In summary, in this study, 15.2% of adults aged 60 and older were found to have mild cognitive impairment. The prevalence of MCI is shown to be higher in females compared to males for the elderly population. Furthermore, GDS scores, smoking, farming, illiteracy, grip strength, hypertension, hyperlipidemia, and stroke may have different associations with MCI in different genders. Additional studies with longer follow-up periods are needed to confirm these associations and to further examine other lifestyle behaviors and diseases that may contribute to MCI. This study may be of great significance to the prevention of MCI and ensuring a healthy aging policy.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Shanghai University of Medicine and Health Sciences. The

patients/participants provided their written informed consent to participate in this study.

Author contributions

YY, QG, and ML contributed to the conception and design of the study. YL wrote the first draft of the manuscript. PH organized the database. XY performed the statistical analysis. XC, FW, JuL, YZ, and ZZ wrote sections of the manuscript. XL, JiL, RL, BW, and CX participate in the survey. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (grant no. 82172552), the Hygiene and Health Commission of Shanghai municipality (2019HP76), the Scientific Research Foundation of SUMHS (SSF-22-03-002), and the Scientific Research Foundation of SUMHS (SSF-21-03-005).

Acknowledgments

We thank Bin Mao and Yue Gu from the Chongming public health center for providing place and organization.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Atkinson, H. H., Rapp, S. R., Williamson, J. D., Lovato, J., Absher, J. R., Gass, M., et al. (2010). The relationship between cognitive function and physical performance in older women: Results from the Women's Health Initiative Memory study. *J. Gerontol. A Biol. Sci. Med. Sci.* 65, 300–306. doi: 10.1093/gerona/glp149

Au, B., Dale-McGrath, S., and Tierney, M. C. (2017). Sex differences in the prevalence and incidence of mild cognitive impairment: A meta-analysis. *Ageing Res. Rev.* 35, 176–199. doi: 10.1016/j.arr.2016.09.005

Chen, X., Han, P., Yu, X., Zhang, Y., Song, P., Liu, Y., et al. (2021). Relationships between sarcopenia, depressive symptoms, and mild cognitive impairment in Chinese community-dwelling older adults. *J. Affect. Disord.* 286, 71–77. doi: 10.1016/j.jad.2021.02.067

Deng, Y., Zhao, S., Cheng, G., Yang, J., Li, B., Xu, K., et al. (2021). The prevalence of mild cognitive impairment among Chinese people: A meta-analysis. *Neuroepidemiology* 55, 79–91. doi: 10.1159/000512597

Dong, L., Briceno, E., Morgenstern, L. B., and Lisabeth, L. D. (2020). Poststroke cognitive outcomes: Sex differences and contributing factors. *J. Am. Heart Assoc.* 9:e16683. doi: 10.1161/JAHA.120.016683

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.* 12, 189–198. doi: 10.1016/0022-3956(75)90026-6

Fu, J., Liu, Q., Du, Y., Zhu, Y., Sun, C., Lin, H., et al. (2020). Age- and sex-specific prevalence and modifiable risk factors of mild cognitive impairment among older adults in china: A population-based observational study. *Front. Aging Neurosci.* 12:578742. doi: 10.3389/fnagi.2020.578742

Ganguli, M., Fu, B., Snitz, B. E., Hughes, T. F., and Chang, C. C. (2013). Mild cognitive impairment: Incidence and vascular risk factors in a population-based cohort. *Neurology* 80, 2112–2120. doi: 10.1212/WNL.0b013e318295d776

Gracia, R. A., Santabarbara, S. J., Lopez, A. R., Tomas, A. C., and Marcos, A. G. (2016). [Occupation and risk of cognitive impairment and dementia in people in over 55 years: A systematic review, Spain]. *Rev. Esp. Salud Publica* 90, e1–e15.

Iadecola, C., Yaffe, K., Biller, J., Bratzke, L. C., Faraci, F. M., Gorelick, P. B., et al. (2016). Impact of hypertension on cognitive function: A scientific statement from the American Heart Association. *Hypertension* 68, e67–e94. doi: 10.1161/HYP.0000000000000053

Jia, L., Du, Y., Chu, L., Zhang, Z., Li, F., Lyu, D., et al. (2020). Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: A cross-sectional study. *Lancet Public Health* 5, e661–e671. doi: 10.1016/S2468-2667(20)30185-7

Jiang, C. Q., Xu, L., Lam, T. H., Zhang, W. S., Liu, B., Lin, J. M., et al. (2009). [Effect of physical activity strength on the diabetes mellitus prevalence in the elderly under the influence of International Physical Activity Questionnaire]. *Zhonghua Liu Xing Bing Xue Za Zhi* 30, 462–465.

Kaiser, M. J., Bauer, J. M., Ramsch, C., Uter, W., Guigoz, Y., Cederholm, T., et al. (2009). Validation of the mini nutritional assessment short-form (MNA-SF): A practical tool for identification of nutritional status. *J. Nutr. Health Aging* 13, 782–788. doi: 10.1007/s12603-009-0214-7

Kim, M., and Park, J. M. (2017). Factors affecting cognitive function according to gender in community-dwelling elderly individuals. *Epidemiol. Health* 39:e2017054. doi: 10.4178/epih.e2017054

Launer, L. J., Andersen, K., Dewey, M. E., Letenneur, L., Ott, A., Amaducci, L. A., et al. (1999). Rates and risk factors for dementia and Alzheimer's disease: Results from EURODEM pooled analyses. EURODEM Incidence Research Group and Work Groups. European Studies of Dementia. *Neurology* 52, 78–84. doi: 10.1212/WNL.52.1.78

Lawton, M. P., and Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *Gerontologist* 9, 179–186. doi: 10.1093/geront/9.3_Part_1.179

Li, J., Wang, Y. J., Zhang, M., Xu, Z. Q., Gao, C. Y., Fang, C. Q., et al. (2011). Vascular risk factors promote conversion from mild cognitive impairment to Alzheimer disease. *Neurology* 76, 1485–1491. doi: 10.1212/WNL.0b013e318217e7a4

Liu, Y., Ma, W., Li, M., Han, P., Cai, M., Wang, F., et al. (2021). Relationship between physical performance and mild cognitive impairment in Chinese community-dwelling older adults. *Clin. Interv. Aging* 16, 119–127. doi: 10.2147/CIA.S288164

Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., et al. (2017). Dementia prevention, intervention, and care. *Lancet* 390, 2673–2734. doi: 10.1016/S0140-6736(17)31363-6

Lv, Y. B., Yin, Z. X., Chei, C. L., Brasher, M. S., Zhang, J., Kraus, V. B., et al. (2016). Serum cholesterol levels within the high normal range are associated with

better cognitive performance among Chinese elderly. *J. Nutr. Health Aging* 20, 280–287. doi: 10.1007/s12603-016-0701-6

Lyu, J., and Kim, H. Y. (2016). Gender-specific incidence and predictors of cognitive impairment among older Koreans: Findings from a 6-year prospective cohort study. *Psychiatry Investig.* 13, 473–479. doi: 10.4306/pi.2016.13.5.473

Ma, W., Zhang, H., Wu, N., Liu, Y., Han, P., Wang, F., et al. (2021). Relationship between obesity-related anthropometric indicators and cognitive function in Chinese suburb-dwelling older adults. *PLoS One* 16:e258922. doi: 10.1371/journal.pone.0258922

Nagamatsu, L. S., Chan, A., Davis, J. C., Beattie, B. L., Graf, P., Voss, M. W., et al. (2013). Physical activity improves verbal and spatial memory in older adults with probable mild cognitive impairment: A 6-month randomized controlled trial. *J. Aging Res.* 2013:861893. doi: 10.1155/2013/861893

National Bureau of Statistics of China (2021). *China statistical yearbook 2019*. Available online at: http://www.stats.gov.cn/tjsj/ndsj/

Ott, A., van Rossum, C. T., van Harskamp, F., van de Mheen, H., Hofman, A., and Breteler, M. M. (1999). Education and the incidence of dementia in a large population-based study: The Rotterdam study. *Neurology* 52, 663–666. doi: 10.1212/WNL.52.3.663

Overton, M., Pihlsgard, M., and Elmstahl, S. (2019). Prevalence and incidence of mild cognitive impairment across subtypes, age, and sex. *Dement. Geriatr. Cogn. Disord.* 47, 219–232. doi: 10.1159/000499763

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., and Kokmen, E. (1999). Mild cognitive impairment: Clinical characterization and outcome. *Arch. Neurol.* 56, 303–308. doi: 10.1001/archneur.56.3.303

Regensteiner, J. G., Golden, S., Huebschmann, A. G., Barrett-Connor, E., Chang, A. Y., Chyun, D., et al. (2015). Sex differences in the cardiovascular consequences of diabetes mellitus: A scientific statement from the American Heart Association. *Circulation* 132, 2424–2447. doi: 10.1161/CIR.0000000000000343

Sharp, E. S., and Gatz, M. (2011). Relationship between education and dementia: An updated systematic review. *Alzheimer Dis. Assoc. Disord.* 25, 289–304. doi: 10.1097/WAD.0b013e318211c83c

Song, M., Wang, Y. M., Wang, R., Xu, S. J., Yu, L. L., Wang, L., et al. (2021). Prevalence and risks of mild cognitive impairment of Chinese community-dwelling women aged above 60 years: A cross-sectional study. *Arch. Womens Ment. Health* 24, 903–911. doi: 10.1007/s00737-021-01137-0

Van der Mussele, S., Fransen, E., Struyfs, H., Luyckx, J., Marien, P., Saerens, J., et al. (2014). Depression in mild cognitive impairment is associated with progression to Alzheimer's disease: A longitudinal study. *J. Alzheimers Dis.* 42, 1239–1250. doi: 10.3233/JAD-140405

Vance, J. E. (2012). Dysregulation of cholesterol balance in the brain: Contribution to neurodegenerative diseases. *Dis. Model. Mech.* 5, 746–755. doi: 10.1242/dmm.010124

Vega, J. N., and Newhouse, P. A. (2014). Mild cognitive impairment: Diagnosis, longitudinal course, and emerging treatments. *Curr. Psychiatry Rep.* 16:490. doi: 10.1007/s11920-014-0490-8

Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., et al. (2004). Mild cognitive impairment–beyond controversies, towards a consensus: Report of the International Working Group on mild cognitive impairment. *J. Intern. Med.* 256, 240–246. doi: 10.1111/j.1365-2796.2004.01380.x

Wu, L., He, Y., Jiang, B., Liu, M., Wang, J., Yang, S., et al. (2016). The association between the prevalence, treatment and control of hypertension and the risk of mild cognitive impairment in an elderly urban population in China. *Hypertens. Res.* 39, 367–375. doi: 10.1038/hr.2015.146

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *J. Psychiatr. Res.* 17, 37–49. doi: 10.1016/0022-3956(82)90033-4

Zhang, M. Y., Katzman, R., Salmon, D., Jin, H., Cai, G. J., Wang, Z. Y., et al. (1990). The prevalence of dementia and Alzheimer's disease in Shanghai, China: Impact of age, gender, and education. *Ann. Neurol.* 27, 428–437. doi: 10.1002/ana.410270412

Zhang, Y., Guan, Y., Shi, Z., Yue, W., Liu, S., Liu, S., et al. (2019). Sex differences in the prevalence of and risk factors for cognitive impairment no dementia among the elderly in a rural area of northern China: A population-based cross-sectional study. *Neuroepidemiology* 52, 25–31. doi: 10.1159/000493141

Zhang, Z. X., Zahner, G. E., Roman, G. C., Liu, X. H., Wu, C. B., Hong, Z., et al. (2006). Socio-demographic variation of dementia subtypes in China: Methodology and results of a prevalence study in Beijing, Chengdu, Shanghai, and Xian. *Neuroepidemiology* 27, 177–187. doi: 10.1159/000096131

frontiers | Frontiers in Aging Neuroscience

TYPE Original Research
PUBLISHED 23 December 2022
DOI 10.3389/fnagi.2022.952145

OPEN ACCESS

EDITED BY
Julie Suhr,
Ohio University,
United States

REVIEWED BY
Ian Brian Malone,
University College London,
United Kingdom
Edmond Teng,
Genentech Inc., United States
Cândida Alves,
Federal University of Maranhão, Brazil

*CORRESPONDENCE
Robert P. Fellows
✉ rfellows@health.ucsd.edu

SPECIALTY SECTION
This article was submitted to Neurocognitive Aging and Behavior, a section of the journal Frontiers in Aging Neuroscience

RECEIVED 24 May 2022
ACCEPTED 09 December 2022
PUBLISHED 23 December 2022

CITATION
Fellows RP, Bangen KJ, Graves LV, Delano-Wood L and Bondi MW (2022) Pathological functional impairment: Neuropsychological correlates of the shared variance between everyday functioning and brain volumetrics. *Front. Aging Neurosci.* 14:952145. doi: 10.3389/fnagi.2022.952145

Pathological functional impairment: Neuropsychological correlates of the shared variance between everyday functioning and brain volumetrics

Robert P. Fellows[1,2]*, Katherine J. Bangen[2,3], Lisa V. Graves[4], Lisa Delano-Wood[1,2] and Mark W. Bondi[1,2]

[1]Psychology Service, VA San Diego Healthcare System, San Diego, CA, United States, [2]Department of Psychiatry, University of California, San Diego, San Diego, CA, United States, [3]Research Service, VA San Diego Healthcare System, San Diego, CA, United States, [4]Department of Psychology, California State University, San Marcos, CA, United States

Objective: Given that several non-cognitive factors can contribute to difficulties with everyday functioning, examining the extent to which cognition is associated with brain-related changes in everyday functioning is critical to accurate characterization of cognitive disorders. In this study, we examined neuropsychological correlates of the shared variance between everyday functioning and pathological indicators of cognitive aging using MRI brain volumetrics.

Participants and methods: Participants were 600 adults aged 55 and older without dementia [432 cognitively normal; 168 mild cognitive impairment (MCI)] from the National Alzheimer's Coordinating Center cohort who underwent neuropsychological testing, informant-rated everyday functioning, and brain MRI scanning at baseline. The shared variance between everyday functioning and brain volumetrics (i.e., hippocampal volume, white matter hyperintensity volume) was extracted using the predicted value from multiple regression. The shared variance was used as an indicator of pathological everyday functional impairment. The residual variance from the regression analysis was used to examine functional reserve.

Results: Larger white matter hyperintensity volumes ($p = 0.002$) and smaller hippocampal volumes ($p < 0.001$) were significantly correlated with worse informant-rated everyday functioning. Among individuals with MCI, worse performances on delayed recall ($p = 0.013$) and category fluency ($p = 0.012$) were significantly correlated with pathological functional impairment in multiple regression analysis. In the cognitively normal group, only worse auditory working memory (i.e., digit span backward; $p = 0.025$) significantly correlated with pathological functioning. Functional reserve was inversely related to anxiety ($p < 0.001$) in the MCI group and was associated with depressive symptoms ($p = 0.003$) and apathy ($p < 0.001$) in the cognitively normal group.

Conclusion: Subtle brain-related everyday functioning difficulties are evident in MCI and track with expected preclinical Alzheimer's disease cognitive phenotypes in this largely amnestic sample. Our findings indicate that functional changes occur early in the disease process and that interventions to target neuropsychiatric symptoms may help to bolster functional reserve in those at risk.

KEYWORDS

instrumental activities of daily living, cognition, white matter hyperintensities, mild cognitive impairment, hippocampus, memory, Alzheimer's disease, attention

Introduction

Older age in adulthood is associated with brain morphological changes (e.g., cortical atrophy, increased white matter disease) and declines in aspects of cognition (Salthouse, 2011; Zhao et al., 2019). The extent to which these factors relate to everyday functioning has important implications for identifying those at risk for pathological cognitive aging. Thresholds for diagnosing neurocognitive disorders rely, in part, on estimating the extent to which cognition contributes to functional ability. One of the primary distinctions between cognitive disorder classifications [e.g., cognitively normal, mild cognitive impairment (MCI), dementia] is the level of independence in instrumental activities of daily living (IADLs). However, several non-cognitive factors (e.g., physical functioning, psychological distress) can influence everyday functioning and complicate clinical and research determinations of brain-mediated functional disability (Vermeulen et al., 2011; Burton et al., 2018).

The term "pathological functional impairment" is used in the current study to represent the overlap between brain morphology consistent with known correlates of cognitive decline (i.e., white matter hyperintensity volume and hippocampal atrophy) and everyday functioning. This approach allows for functional difficulties related to brain volumetrics to be examined as both a continuous variable across the cognitive aging spectrum as well as within and between discrete diagnostic classifications (e.g., cognitively normal, MCI). The term "pathological functional impairment" is not intended to provide a static definition, but rather it is intended to represent a dynamic concept that can be applied to the study of disease in different populations and evolve and change.

In a systematic review of 20 studies examining correlates of everyday functioning in older adults, neuropsychological performance and brain morphology (hippocampal atrophy and white matter changes, in particular) were both unique and overlapping predictors of IADLs (Overdorp et al., 2016). These findings suggest a differential influence of structural brain changes on cognition and functional decline. Thus, examining cognitive correlates of the shared variance between brain markers of pathological aging and everyday functioning may help improve the ecological validity of neuropsychological tests, such that tests associated with the shared variance may be surrogates of or metrics reflecting subtle functional difficulties corresponding to early neurodegenerative changes. In the current study, we used regression analysis to isolate the shared variance between MRI-derived brain variables and an informant report of everyday functioning.

White matter hyperintensities (WMHs), shown as areas of high signal in white matter on T2-weighted magnetic resonance imaging (MRI), are indicators of cerebral small vessel disease and are associated with cognitive decline (Decarli et al., 2001; Brickman et al., 2008; Carmichael et al., 2010; Hu et al., 2021). WMHs have also been associated with a faster rate of everyday functioning decline in older adults without dementia (Puzo et al., 2019; Bangen et al., 2020) and progression from cognitively normal to MCI (DeCarli et al., 2001; Bangen et al., 2018). Hippocampal atrophy is another indicator of neurodegeneration and cognitive decline often observed in Alzheimer's disease (Barnes et al., 2009; Sabuncu et al., 2011). Studies have shown that smaller hippocampal volumes in cognitively normal individuals are associated with incident MCI (Apostolova and Cummings, 2008; Bangen et al., 2018). Given that changes in these brain structures are often evident long before clinical detection, identifying cognitive correlates of the shared variance between brain volumetric indicators of pathological processes and everyday functioning may have important implications for improving the early detection of neurodegenerative disease.

In addition to identifying neuropsychological markers of pathological everyday functioning, it is important to examine non-cognitive factors that may contribute to daily functioning in order to identify other potentially modifiable risk factors for functional decline. Studies have examined this concept in the context of "functional reserve" or the variance in everyday functioning that is not explained by demographics, neuropsychological performance, or brain volumetrics (Berezuk et al., 2017; Kraal et al., 2021). A recent study examining predictors of functional reserve found that physical function, informant-rated apathy, and informant-rated depression were significantly associated with the residual variance in everyday functioning in a

large sample of older adults with normal cognition, MCI, and dementia (Kraal et al., 2021). However, a closer investigation of the way in which correlates of functional reserve may differ as a function of diagnostic status (i.e., cognitively normal vs. MCI) is needed.

In the present study, we aimed (1) to identify the shared variance between brain volumetric indicators of cognitive aging (i.e., WMH and hippocampal volumes) and everyday functioning, (2) to examine demographic, genetic (i.e., APOE ε4 status), and neuropsychological factors associated with the shared variance between brain volumetrics and everyday functioning (i.e., pathological functioning) in individuals with normal cognition and those with MCI, and (3) to identify correlates of the unique or residual variance in everyday functioning (i.e., functional reserve) not accounted for by demographic, neuropsychological, or brain variables. Consistent with these aims, we hypothesized that (1) larger WMH volumes and smaller hippocampal volumes would account for a significant amount of variance in everyday functioning, (2) APOE ε4 carrier status and poorer neuropsychological test performance would be associated with pathological functioning, particularly in the MCI group, and (3) neuropsychiatric symptoms would account for a significant amount of variance in functional reserve.

Materials and methods

Data used in the current study were derived from the National Alzheimer's Coordinating Center (NACC) Uniform Data Set (UDS)[1] up through the September 2020 data freeze and includes data from 14 National Institute on Aging (NIA)-funded Alzheimer's Disease Research Centers (ADRC). These ADRCs contributed neuropsychological, medical, neuroimaging, and functional data to the NACC-UDS. All ADRCs received approval from their local Institutional Review Boards (IRB) and obtained informed consent from study participants. As determined by the University of Washington Human Subjects Division, the NACC database itself is exempt from IRB review and approval because it does not involve human subjects, as defined by federal and state regulations. In the current study, data from baseline UDS visits conducted between 2005 and 2015 were used in the analyses.

Participants were 600 individuals (168 MCI; 432 cognitively normal) age 55 and older with baseline neuropsychological data, informant-rated everyday functioning, APOE ε4 genotyping, as well as MRI scan data with volumes within 6 months of the baseline UDS visit. Exclusion criteria for the current study were a diagnosis of dementia or other neurological condition (stroke, TBI, seizures, Parkinson's disease), less than 6 years of formal education, or a Geriatric Depression Scale score of 6 or higher.

1 http://www.alz.washington.edu

NACC diagnostic classifications

Classification of cognitive status (i.e., cognitively normal vs. MCI) was determined by clinician evaluation or by consensus at each ADRC using all available data including demographics, medical history, medical evaluations, neuropsychological test scores, Clinical Dementia Rating CDR® Dementia Staging Instrument (CDR) score, and informant-rated everyday functioning and emotional functioning (Morris et al., 2006; Beekly et al., 2007; Weintraub et al., 2009). Data included in the current study are from individuals who were classified as either cognitively normal or MCI at the baseline visit.

Assessment of everyday functioning

The Functional Activities Questionnaire (FAQ) is a 10-item questionnaire (Pfeffer et al., 1982) used to assess everyday functioning including (1) writing checks, paying bills, or balancing a checkbook; (2) assembling tax records, business affairs, or other papers; (3) shopping alone for clothes, household necessities, or groceries; (4) playing a game of skill such as bridge or chess, or working on a hobby; (5) heating water, making a cup of coffee, or turning off the stove; (6) preparing a balanced meal; (7) keeping track of current events; (8) paying attention to and understanding a TV program, book, or magazine; (9) remembering appointments, family occasions, holidays, medications; and (10) traveling out of the neighborhood, driving, or arranging to take public transportation. Each item is rated on a scale from 0 (normal) to 3 (dependent) with a possible range of 0 to 30 with higher total scores indicating greater dependence. If the activity queried was never completed by the participant, it was coded as "not applicable." For the purposes of the current study, "not applicable" responses were coded as 0 (Teng et al., 2010). Informant ratings of everyday functioning (i.e., FAQ) were completed in interviews with study partners (Weintraub et al., 2009). Mean FAQ scores were used in the analyses.

Neuropsychological assessment

A standardized battery of neuropsychological tests was administered at each NACC study visit. Details of the NACC-UDS neuropsychological battery have been described previously (Weintraub et al., 2009). The battery included the Wechsler Memory Scale (WMS) Logical Memory, Story A, Immediate and Delayed Recall; Trail Making Test Part A and B; Vegetable Fluency and Animal Fluency; Boston Naming Test (BNT); Digit Span Forward and Backward; and Wechsler Adult Intelligence Scale (WAIS) Digit Symbol. Raw scores were used in all analyses in the present study.

Genotyping

APOE ε4 status (carrier vs. non-carrier) was derived from the NACC genetic dataset. Participants were classified as APOE ε4 positive if they had at least one APOE ε4 allele.

MRI brain volumetrics

A subset of Alzheimer's Disease Centers (ADCs) voluntarily submits MR images to NACC. Imaging data collection and acquisition protocols vary by ADC. The majority of scans were either 1.5 T ($n = 241$) or 3 T ($n = 273$) with the remainder reportedly varying in field strength across images or not reported. MRI manufacturers included GE ($n = 452$), Siemens ($n = 97$), and Philips ($n = 44$) or not reported ($n = 7$). Volumetric MRI quantification was performed by the Imaging of Dementia & Aging (IDeA) Lab (Director: Charles DeCarli, MD; University of California, Davis).[2] Total WMH volume (in cm^3) was calculated from T2-weighted fluid-attenuated inversion recovery (FLAIR) images. Details regarding FLAIR image acquisition have been described in a previous publication (Alosco et al., 2018) and processing was similar to that used in ADNI2 (detailed description of processing can be found on the NACC website).[3] Briefly, the FLAIR image was co-registered to the high-resolution T1-weighted image and inhomogeneity correction was applied to both images (DeCarli et al., 1996). Images were non-linearly transformed into a common template atlas. WMH segmentation was based on a modified Bayesian approach combining image likelihood estimates, spatial priors, and tissue class constraints. WMH prior probability maps were derived based on more than 700 participants using semi-automatic detection of WMH and subsequent manual editing. Likelihood estimates of the native image were calculated *via* histogram segmentation and thresholding. Segmentation was initially performed in standard space resulting in probability likelihood values of WMH at each white matter voxel. These probabilities were then thresholded at >3.5 SD above the mean to create a binary WMH mask (DeCarli et al., 1999). The segmented WMH masks were then back-transformed into native space for volume calculation. Voxels labeled as WMH were summed and multiplied by voxel dimensions to obtain volumes and reported in units of cm^3. Hippocampal and total cerebral volumes were determined on the T1-weighted image using automated procedures performed by the IDeA lab. The hippocampal segmentation method uses a standard atlas-based diffeomorphic approach (Vercauteren et al., 2007) with minor label refinement modifications. This approach was further modified to include the European Alzheimer's Disease Consortium (EADC)-Alzheimer's Disease Neuroimaging Initiative (ADNI) harmonized hippocampal masks to assure standardization across cohorts using the following procedures: (1) Subject image pre-processing with extraction of intracranial cavity, non-uniformity correction, tissue classification; (2) Atlas Registration of all EADC-ADNI hippocampal masks (Frisoni et al., 2013; Boccardi et al., 2015); (3) atlas fusion utilizing Multi-Atlas Label Fusion (Wang et al., 2012); and (4) Intensity-based label refinement. Total hippocampal volume was the sum of right and left hippocampi volumes. Total brain volume was calculated by summing total gray and white matter volumes.

Neuropsychiatric symptoms

Informant ratings of psychiatric symptoms were obtained *via* structured clinical interviews using the Neuropsychiatric Inventory Questionnaire (NPI-Q; Kaufer et al., 2000). The severity of symptoms ranges from 0 (no symptoms) to 3 (severe symptoms). Single-item ratings of apathy, anxiety and depression were selected for inclusion in the current study based on previous research showing associations with everyday functioning (Hwang et al., 2004; Rog et al., 2014; Kraal et al., 2021).

Statistical analyses

To establish a pathological functioning score representing the shared variance between brain volumetric indicators and everyday functioning, using a hierarchical regression analysis, mean FAQ scores (log-transformed to improve distribution normality due to positive skew) were regressed on total hippocampal volume and total white matter hyperintensity volume (log-transformed to improve distribution normality due to positive skew), adjusting for log-transformed total intracranial volume (ICV). Both the standardized predicted (i.e., pathological functioning variable) and standardized residual (i.e., functional reserve variable) values were used as dependent variables in subsequent analyses. The functional reserve variable was reverse signed so that higher values indicate greater functional reserve.

To identify cognitive correlates of pathological functioning, hierarchical regression analyses were conducted with the pathological functioning value regressed on age, sex, race/ethnicity, and APOE ε4 status (carrier vs. non-carrier) in block 1, and neuropsychological test scores in block 2. Separate regression analyses were used to examine cognitive correlates of pathological functioning in the cognitively normal group and those classified as MCI.

To identify neuropsychiatric correlates of functional reserve, the residual variance value was regressed on NPI-Q apathy, depression, and anxiety scores (in block 3), adjusting for demographics and APOE ε4 status (in block 1), and neuropsychological test variables (in block 2). These analyses were stratified by cognitive status (cognitively normal, MCI). Possible multicollinearity effects were examined and determined to be within an acceptable range (i.e., all VIFs <4).

2 http://idealab.ucdavis.edu/

3 https://files.alz.washington.edu/documentation/adni-proto.pdf

10.3389/fnagi.2022.952145

TABLE 1 Participant characteristics.

	Total sample (N=600)		Cognitively normal (n=432)		MCI (n=168)			Effect size	
	M	SD	*M*	SD	*M*	SD	*F* or X^2	Hedge's *g* or phi	*p*
Demographics									
Age	71.32	8.67	69.75	8.63	75.36	7.40	55.02	0.674	<0.001
Sex (% female)	58.5		62.5		48.2		10.16	0.130	0.002
Education	15.50	3.10	15.63	3.04	15.15	3.26	2.93	0.155	0.087
Race/ethnicity (% NHW)	91.2		90.5		92.9		0.828	0.037	0.425
Genetics									
APOE ε4 (% positive)	36.7		32.9		46.4		9.57	0.126	0.002
Neuropsychological									
LM I – Immediate Recall	11.46	4.53	12.93	3.94	7.68	3.68	222.24	1.354	<0.001
LM II – Delayed Recall	9.81	5.11	11.69	4.23	4.96	3.81	322.94	1.632	<0.001
Digit Span Forward	8.29	2.03	8.38	2.07	8.07	1.91	2.79	0.152	0.095
Digit Span Backward	6.47	2.21	6.74	2.23	5.76	2.01	24.51	0.450	<0.001
Animal Fluency	19.50	6.09	21.09	5.77	15.42	4.87	127.43	1.025	<0.001
Vegetable Fluency	13.69	4.69	15.02	4.40	10.27	3.53	155.76	1.133	<0.001
Digit Symbol	45.66	12.81	49.12	11.71	36.76	11.15	138.23	1.068	<0.001
Boston Naming Test	26.59	4.06	27.67	2.86	23.82	5.22	132.88	1.047	<0.001
Trails A	36.22	18.83	32.05	13.70	46.95	25.05	86.44	0.844	<0.001
Trails B	102.57	62.31	83.94	42.78	150.49	77.42	179.03	1.215	<0.001
Everyday functioning									
FAQ	0.15	0.36	0.05	0.20	0.42	0.50	168.64	1.179	<0.001
Brain volumetrics									
Total Brain volume	0.71	0.16	0.72	0.16	0.69	0.16	3.78	0.344	0.052
Total WMH volume	1.11	2.16	0.82	2.18	1.83	1.92	28.17	0.482	<0.001
Hippocampal volume	6.26	0.81	6.44	0.71	5.81	0.87	82.63	0.764	<0.001
Neuropsychiatric									
NPI-Q Depression	0.20	0.52	0.14	0.45	0.34	0.63	19.02	0.397	<0.001
NPI-Q Anxiety	0.17	0.48	0.09	0.35	0.34	0.68	34.97	0.538	<0.001
NPI-Q Apathy	0.11	0.38	0.05	0.27	0.23	0.55	29.63	0.495	<0.001

Effect sizes are in terms of phi for categorical variables and Hedge's g for continuous variables.

Results

Participant characteristics and descriptive statistics

Sample characteristics are shown in Table 1. As expected, the MCI group was older, more likely to be APOE ε4 positive, and had lower neuropsychological test scores, higher mean FAQ scores, smaller hippocampal volumes, and greater WMH volume compared to cognitively normal participants. The MCI group also had higher informant-rated depression, anxiety, and apathy severity.

Brain volumetric correlates of everyday functioning

In the hierarchical multiple regression analysis, ICV was entered in the first block and WMH volume and hippocampal volume variables were entered into the second block. The results indicated that the first block was not statistically significant [$R^2 = 0.000$, $F(1, 598) = 0.078$, $p = 0.781$]. Results of the second block indicated that the model was statistically significant [$\Delta R^2 = 0.089$, $\Delta F(2, 596) = 28.938$, $p < 0.001$]. Greater total WMH volumes ($B = 0.024$, SE = 0.006, $\beta = 0.229$, $t = 4.036$, $p < 0.001$) and smaller hippocampal volumes ($B = -0.056$, SE = 0.012, $\beta = -0.200$,

TABLE 2 Hierarchical regression model of cognitive correlates of pathological functioning in MCI group.

Model					t	p
		B	SE	β		
1	(Constant)	−3.423	0.978		−3.499	<0.001
	Age	0.044	0.009	0.346	4.755	<0.001
	Sex	0.554	0.138	0.297	4.023	<0.001
	Education	−0.021	0.021	−0.074	−1.002	0.318
	Race/ethnicity	0.077	0.259	0.021	0.297	0.766
	E4 status	0.382	0.135	0.204	2.827	0.005
	$R^2 = 0.233$, $F(5, 162) = 9.845$, $p < 0.001$					
2	(Constant)	−2.317	1.115		−2.079	0.039
	Age	0.041	0.009	0.323	4.425	<0.001
	Sex	0.632	0.145	0.339	4.354	<0.001
	Education	−0.026	0.022	−0.090	−1.168	0.245
	Race/ethnicity	−0.018	0.254	−0.005	−0.072	0.943
	E4 status	0.250	0.132	0.134	1.901	0.059
	LM I - Immediate Recall	0.037	0.028	0.145	1.307	0.193
	LM II - Delayed Recall	−0.068	0.027	−0.275	−2.474	0.014
	Digit Span Forward	−0.056	0.039	−0.115	−1.439	0.152
	Digit Span Backward	0.036	0.039	0.078	0.929	0.354
	Animal Fluency	−0.002	0.016	−0.011	−0.127	0.899
	Vegetable Fluency	−0.052	0.022	−0.198	−2.390	0.018
	Digit Symbol	0.016	0.008	0.193	2.060	0.041
	Boston Naming Test	−0.028	0.014	−0.157	−1.962	0.052
	Trails A	0.006	0.003	0.158	1.818	0.071
	Trails B	−0.001	0.001	−0.059	−0.658	0.511
	$\Delta R^2 = 0.146$, $\Delta F(10, 152) = 3.561$, $p < 0.001$					

$N = 168$.

$t = -4.783$, $p < 0.001$) were significantly associated with higher FAQ scores. The standardized predicted variance value was used as the pathological functioning score, with higher values indicating greater brain-related difficulties with everyday functioning. The residual variance, representing everyday functioning unrelated to normalized total brain volume, hippocampal volume, and WMH volume, was used in subsequent functional reserve analyses.

Given that only 460 of the 600 participants had complete data for all 10 FAQ items, an exploratory regression analysis was conducted that included only those with complete FAQ data to examine whether the results differed if those with missing data were excluded. Consistent with the analysis of the whole sample, block 1 was non-significant [$R^2 = 0.000$, $F(1, 458) = 0.212$, $p = 0.645$] and block 2 was statistically significant [$\Delta R^2 = 0.083$, $\Delta F(2, 456) = 20.697$, $p < 0.001$] with WMH volumes [$B = 0.015$, SE = 0.006, $\beta = 0.172$, $t = 2.632$, $p = 0.009$] and hippocampal volumes ($B = -0.054$, SE = 0.012, $\beta = -0.223$, $t = -4.637$, $p < 0.001$) associated with mean FAQ scores. Therefore, data from the entire sample ($n = 600$) was used in the analyses.

Neuropsychological correlates of pathological functioning

In the hierarchical regression analysis for the MCI group (see Table 2), demographic characteristics and APOE ε4 status were entered into the first block and neuropsychological variables were entered into the second block. Results of the model in the first block were statistically significant [$R^2 = 0.212$, $F(5, 162) = 8.729$, $p < 0.001$]. Older age, female sex, and APOE ε4 positivity were associated with a higher pathological functioning score. The model in the second block was statistically significant [$\Delta R^2 = 0.154$, $\Delta F(10, 152) = 3.680$, $p < 0.001$], and performances on delayed recall (LMII) and vegetable fluency were negatively, and digit symbol substitution test positively associated with the pathological functioning score. A trend-level association was observed for BNT ($p = 0.050$).

In the cognitively normal group (see Table 3), results of the hierarchical regression revealed that the first block was significant [$R^2 = 0.379$, $F(5, 426) = 51.892$, $p < 0.001$] and that older age, female sex, less education, and non-NHW race/ethnicity was associated with greater pathological functional impairment. The model in the second block was statistically significant [$\Delta R^2 = 0.032$, $\Delta F(10, 416) = 2.274$, $p = 0.013$], with only Digit Span Backwards as a significant predictor.

Neuropsychiatric correlates of functional reserve

The residual variance from the regression analysis with brain volumetrics predicting FAQ scores was used as the dependent variable in the functional reserve analyses. As with the previous analyses, demographics, APOE ε4 status, and neuropsychological variables were entered into the first blocks (block 1 and block 2) and then NPI-Q depression, anxiety, and apathy were entered into the final block (block 3) to examine the unique association of emotional factors in everyday functioning.

TABLE 3 Hierarchical regression model of cognitive correlates of pathological functioning in cognitively normal group.

Model					*t*	*p*
		B	SE	*β*		
1	(Constant)	−5.027	0.449		−11.185	<0.001
	Age	0.060	0.004	0.570	14.812	<0.001
	Sex	0.442	0.072	0.234	6.137	<0.001
	Education	−0.031	0.012	−0.104	−2.577	0.010
	Race/ethnicity	0.275	0.125	0.088	2.209	0.028
	E4 status	0.122	0.074	0.063	1.656	0.099
	$R^2 = 0.399$, $F(5, 426) = 56.553$, $p < 0.001$					
2	(Constant)	−3.244	0.723		−4.485	<0.001
	Age	0.050	0.005	0.474	10.092	<0.001
	Sex	0.458	0.080	0.243	5.762	<0.001
	Education	−0.007	0.013	−0.025	−0.567	0.571
	Race/ethnicity	0.115	0.129	0.037	0.892	0.373
	E4 status	0.078	0.073	0.040	1.068	0.286
	LM I - Immediate Recall	0.016	0.016	0.070	0.991	0.322
	LM II - Delayed Recall	−0.023	0.015	−0.108	−1.540	0.124
	Digit Span Forward	−0.005	0.020	−0.011	−0.227	0.821
	Digit Span Backward	−0.047	0.019	−0.115	−2.434	0.015
	Animal Fluency	−0.006	0.008	−0.039	−0.755	0.451
	Vegetable Fluency	−0.004	0.010	−0.017	−0.347	0.729
	Digit Symbol	−0.004	0.005	−0.046	−0.770	0.442
	Boston Naming Test	−0.019	0.015	−0.059	−1.260	0.208
	Trails A	−0.001	0.003	−0.011	−0.223	0.824
	Trails B	0.000	0.001	0.019	0.319	0.750
	$\Delta R^2 = 0.033$, $\Delta F(10, 416) = 2.419$, $p = 0.008$					

$N = 432$.

In the MCI group ($n = 167$), block 1 was significant [$R^2 = 0.085$, $F(5, 161) = 3.009$, $p = 0.013$], with male sex as the only statistically significant variable [$B = 0.709$, SE = 0.223, $\beta = 0.256$, $t = 3.171$, $p = 0.002$]. Block 2, with neuropsychological variables, was not statistically significant [$\Delta R^2 = 0.091$, $\Delta F(10, 151) = 1.662$, $p = 0.095$]. Block 3 was statistically significant [$\Delta R^2 = 0.118$, $\Delta F(3, 148) = 8.230$, $p < 0.001$], and only anxiety severity was associated with the residual variance score ($B = -0.692$, SE = 0.164, $\beta = -0.342$, $t = -4.220$, $p < 0.001$), such that greater anxiety was associated with lower functional reserve.

In the cognitively normal group, block 1 was significant [$R^2 = 0.048$, $F(5, 426) = 4.250$, $p < 0.001$], and older age ($B = 0.010$, SE = 0.004, $\beta = 0.142$, $t = 2.930$, $p = 0.004$), male sex ($B = 0.185$, SE = 0.062, $\beta = 0.143$, $t = 2.986$, $p = 0.003$), and APOE ε4 negativity ($B = -0.129$, SE = 0.064, $\beta = -0.097$, $t = -2.019$, $p = 0.044$) were associated with the residual variance score. Neuropsychological variables did not account for a significant amount of variance in block 2 [$\Delta R^2 = 0.036$, $\Delta F(10, 416) = 1.642$, $p = 0.092$]. Results in block 3 were significant [$\Delta R^2 = 0.087$, $\Delta F(3, 413) = 14.412$, $p < 0.001$], with greater apathy severity ($B = -0.462$, SE = 0.106, $\beta = -0.207$, $t = -4.353$, $p < 0.001$) and greater depression severity ($B = -0.208$, SE = 0.068, $\beta = -0.151$, $t = -3.033$, $p = 0.003$) associated with lower functional reserve.

Discussion

We found that pathological everyday functioning, defined as the shared variance between informant-rated FAQ scores and MRI brain variables (i.e., WMH and hippocampal volumes), was differentially associated with cognition as a function of cognitive status (MCI vs. cognitively normal). In the MCI group, cognitive variables accounted for 15.5% of the total variance in pathological functioning compared to 3.2% in the cognitively normal group. These findings suggest that subtle brain-related everyday functional difficulties are evident in individuals without dementia and correspond to the level of cognitive functioning observed on neuropsychological measures. Previous research has shown that mild functional impairments are present in MCI (Burton et al., 2009; Bangen et al., 2010; Teng et al., 2010) as well as cognitively normal individuals (Farias et al., 2017). Our results support and expand these findings by demonstrating aspects of cognitive functioning in MCI that are associated with everyday functioning that is uniquely related to MRI brain variables indicative of pathological processes (e.g., atrophy, white matter disease). This approach provides a clearer link between cognition and everyday functioning that is not affected by potentially confounding factors such as peripherally-mediated physical functioning and/or emotional distress.

Our secondary aim was to examine neuropsychiatric correlates of functional reserve, the unique variance in everyday functioning not accounted for by demographics, cognition, or brain volumetrics. Consistent with prior research using samples of adults with normal cognition, MCI, and dementia (Kraal et al., 2021), we found that greater neuropsychiatric symptomatology was associated with lower functional reserve. We extended prior work by stratifying our sample by cognitive status (cognitively normal, MCI) and found that in the MCI group, only anxiety severity, but not depression or apathy, was associated with lower

functional reserve in the multivariable model. The lack of association of depression and apathy with functional reserve in the MCI group was unexpected, particularly because these symptoms are prevalent in MCI and previous studies have shown an association with everyday functioning (Zahodne and Tremont, 2013; Rog et al., 2014) in MCI. However, these prior studies did not examine the extent to which neuropsychiatric factors are associated with functional reserve, which may account for the discrepant findings.

Given the evidence that both WMH and hippocampal volume predict incident AD in older adults without dementia (Brickman et al., 2015), and that acceleration of WMH accumulation precedes clinically detected cognitive impairment (Silbert et al., 2008, 2009, 2012), we were interested in examining whether these variables were associated with subtle functional difficulties in older adults without dementia. Consistent with previous studies, we found that greater WMH and smaller hippocampal volumes accounted for a significant amount of variance in informant-rated everyday functioning among older adults without dementia (DeCarli et al., 2001; Apostolova and Cummings, 2008; Bangen et al., 2018). By combining these variables into a single metric (i.e., pathological functioning), we could be inclusive of common early etiologic indicators of cognitive decline, specifically, cerebral small vessel disease and neurodegeneration consistent with AD pathogenesis. This approach is advantageous because it restricts informant–rated functional ability into a variable that fully overlaps with MRI indicators of brain pathology and, thus, is likely to be more sensitive to subtle cognitive changes.

In the MCI group, the pattern of association between cognitive variables and pathological functioning was consistent with an AD process such that lower verbal memory and semantic fluency scores predicted pathological everyday functioning in the cross-sectional regression analysis. This finding supports the ecological validity of these measures in detecting early IADL changes as they were significantly associated with brain-related aspects of everyday functioning. In contrast, a measure of processing speed (i.e., digit symbol) was associated with pathological functioning, although it was in the opposite direction than expected such that faster processing speed was associated with greater pathological functioning. A similar trend ($p = 0.064$) was observed for Trails A performance. The reason for this counterintuitive direction is unclear and requires further investigation. Follow-up partial correlations used to examine this unexpected result revealed that the association between processing speed and pathological functioning was not statistically significant (see Appendix). The lack of significant association between executive functioning (as measured by Trails B performance) and everyday functioning in the current study contrasts with a meta-analytic study (McAlister et al., 2016) showing Trails B to be the strongest cognitive predictor of functional status in MCI. Although WMH volumes have been linked to poorer executive functioning in older adults (e.g., Boutzoukas et al., 2021), in the current study we did not examine other brain variables associated with executive functioning such as frontal lobe structures, which may account for the null association we found between Trails B and pathological everyday functioning. It is also likely that this difference is due to the large representation of amnestic MCI (84.5%) subtype in the composition of the sample used in our study.

Prior work has shown that early alterations in attentional control are sensitive to AD-related biomarkers in cognitively healthy individuals (Aschenbrenner et al., 2015, 2020; McKay et al., 2022). Consistent with previous research, in the current study working memory (Digit Span Backward) was the only cognitive predictor of pathological functioning in the cognitively normal group. This finding suggests that subtle cognitive changes in cognitively normal individuals may correspond to very early mild brain-related functional changes. However, given the relatively small effect, further research is needed to verify the reliability of this association.

Although cognitive variables accounted for a significant portion of pathological functioning variance across diagnostic groups, the contribution of demographic characteristics and APOE ε4 carrier status to the regression models were notable. In the initial models, these variables accounted for approximately 21% and 38% of the variance in pathological functioning in MCI and cognitively normal individuals, respectively. In the MCI group, older age, female sex, and APOE ε4 status remained significant correlates of pathological functioning after the inclusion of neuropsychological variables in the model. Among cognitively normal participants, older age, female sex, less education, and self-identifying as not non-Hispanic white were significant correlates in the initial regression model, but only age and sex remained significant after including cognitive variables. These findings raise the possibility that traditional neuropsychological tests are not capturing the full spectrum of age-related cognitive changes involved in complex activities of daily living. Although the measures used in the current study have been well-validated and spanned multiple cognitive domains, some aspects of cognition relevant to cognitive aging were not examined, notably, several aspects of memory such as prospective memory, temporal order memory, and visual memory. Indeed, prior work has shown that performance on novel memory paradigms is associated with self-reported, informant-reported and directly observed measures of everyday functioning (Fellows and Schmitter-Edgecombe, 2018). It is also possible that physical capacity may mediate or moderate the association between age and pathological functioning, independent from cognition. Studies have shown that WMHs (Willey et al., 2013), and smaller medial temporal regions (Rosano et al., 2012; Ezzati et al., 2015) are associated with slowed gait speed. Reduced mobility has been shown to mediate the association between age and the ability to perform daily activities that require physical capacity (Fellows and Schmitter-Edgecombe, 2019). These studies suggest pathways between age-related brain changes and mobility that may influence everyday functioning independently from cognition. Yet, given the shared etiology of mobility and cognitive changes *via* global WMH and medial temporal atrophy, these factors may exert an additive effect on everyday functioning.

The NACC data used in this study include a large well-characterized group of participants with neuropsychological and neuroimaging data. However, the participants are generally highly educated and mostly non-Hispanic white, which limits the ability to generalize findings to more diverse groups that represent nearly 40% of the U.S. adult population (U.S. Census, 2020). Although the NACC dataset overall has greater racial/ethnic diversity than other large prospective cohort studies (Kiselica and Benge, 2021), restricting inclusion in the current study to only individuals who had neuropsychological, MRI, and everyday functioning data resulted in a sample that was >90% non-Hispanic white. Further, Gleason et al. (2019) demonstrated that NACC study enrollment factors, such as referral source (i.e., Black participants being over-sampled in community-based recruitment and under-sampled in clinic-based recruitment), introduce further racial/ethnic bias into an already under-representative sample of Black participants. In the current study, race/ethnicity was significantly associated with pathological functioning in the cognitively normal group but the correlation was attenuated to a non-significant level after the inclusion of cognitive variables in the regression model. This finding may suggest a possible interactive effect between race/ethnicity and cognition on pathological functioning, however, due to the small sample size and bias in sampling, further examination of this effect could not be examined. Greater diversity in large research cohorts is needed to better understand the complex factors contributing to racial and ethnic differences in cognitive aging.

The cross-sectional and observational design of the analyses used in the current study precludes the interpretation of causality. Additionally, by design, participants included in this study had only minimal informant-rated functional difficulties and minimal depressive symptoms. Future research examining the full spectrum of cognitive aging without exclusion of individuals with more severe emotional distress is needed to further elucidate the continuum of pathological functioning and functional reserve.

In summary, we found that lower hippocampal volumes and greater total WMH volumes are independently associated with increased functional difficulties in adults without dementia. Our findings indicate that subtle brain-related everyday functioning difficulties are evident in MCI and correspond with the preclinical Alzheimer's disease cognitive phenotype (lower verbal memory and category fluency) which highlights the importance of neuropsychological assessment in the early detection of pathological cognitive aging. We also found that even relatively low levels of informant-rated neuropsychiatric symptoms were associated with lower functional reserve and may be targets for early therapeutic intervention.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: https://naccdata.org/.

Author contributions

RF designed the study, analyzed and interpreted the data, and drafted the manuscript. KB, LG, LD-W, and MB contributed critical intellectual content to manuscript revisions. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by Grants R01 AG049810 and R01 AG063782. The NACC database is funded by NIA/NIH Grant U24 AG072122. NACC data are contributed by the NIA-funded ADCs: P50 AG005131 (PI James Brewer, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P50 AG005138 (PI Mary Sano, PhD), P50 AG005142 (PI Helena Chui, MD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005681 (PI John Morris, MD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG008051 (PI Thomas Wisniewski, MD), P50 AG008702 (PI Scott Small, MD), P30 AG010124 (PI John Trojanowski, MD, PhD), P30 AG010129 (PI Charles DeCarli, MD), P30 AG010133 (PI Andrew Saykin, PsyD), P30 AG010161 (PI David Bennett, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG013846 (PI Neil Kowall, MD), P30 AG013854 (PI Robert Vassar, PhD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P30 AG019610 (PI Eric Reiman, MD), P50 AG023501 (PI Bruce Miller, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P30 AG028383 (PI Linda Van Eldik, PhD), P50 AG033514 (PI Sanjay Asthana, MD, FRCP), P30 AG035982 (PI Russell Swerdlow, MD), P50 AG047266 (PI Todd Golde, MD, PhD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG049638 (PI Suzanne Craft, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG066546 (PI Sudha Seshadri, MD), P20 AG068024 (PI Erik Roberson, MD, PhD), P20 AG068053 (PI Marwan Sabbagh, MD), P20 AG068077 (PI Gary Rosenberg, MD), P20 AG068082 (PI Angela Jefferson, PhD), P30 AG072958 (PI Heather Whitson, MD), and P30 AG072959 (PI James Leverenz, MD).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Alosco, M. L., Sugarman, M. A., Besser, L. M., Tripodis, Y., Martin, B., Palmisano, J. N., et al. (2018). A clinicopathological investigation of white matter hyperintensities and Alzheimer's disease neuropathology. *J. Alzheimers Dis.* 63, 1347–1360. doi: 10.3233/JAD-180017

Apostolova, L. G., and Cummings, J. L. (2008). Neuropsychiatric manifestations in mild cognitive impairment: a systematic review of the literature. *Dement. Geriatr. Cogn. Disord.* 25, 115–126. doi: 10.1159/000112509

Aschenbrenner, A. J., Balota, D. A., Fagan, A. M., Duchek, J. M., Benzinger, T. L., and Morris, J. C. (2015). Alzheimer disease cerebrospinal fluid biomarkers moderate baseline differences and predict longitudinal change in attentional control and episodic memory composites in the adult children study. *J. Int. Neuropsychol. Soc.* 21, 573–583. doi: 10.1017/S1355617715000776

Aschenbrenner, A. J., Gordon, B. A., Fagan, A. M., Schindler, S. E., Balota, D. A., Morris, J. C., et al. (2020). Neurofilament light predicts decline in attention but not episodic memory in preclinical Alzheimer's disease. *J. Alzheimers Dis.* 74, 1119–1129. doi: 10.3233/JAD-200018

Bangen, K. J., Jak, A. J., Schiehser, D. M., Delano-Wood, L., Tuminello, E., Han, S. D., et al. (2010). Complex activities of daily living vary by mild cognitive impairment subtype. *J. Int. Neuropsychol. Soc.* 16, 630–639. doi: 10.1017/S1355617710000330

Bangen, K. J., Preis, S. R., Delano-Wood, L., Wolf, P. A., Libon, D. J., Bondi, M. W., et al. (2018). Baseline white matter hyperintensities and hippocampal volume are associated with conversion from normal cognition to mild cognitive impairment in the Framingham offspring study. *Alzheimer Dis. Assoc. Disord.* 32, 50–56. doi: 10.1097/WAD.0000000000000215

Bangen, K. J., Thomas, K. R., Weigand, A. J., Sanchez, D. L., Delano-Wood, L., Edmonds, E. C., et al. (2020). Pattern of regional white matter hyperintensity volume in mild cognitive impairment subtypes and associations with decline in daily functioning. *Neurobiol. Aging* 86, 134–142. doi: 10.1016/j.neurobiolaging.2019.10.016

Barnes, J., Bartlett, J. W., van de Pol, L. A., Loy, C. T., Scahill, R. I., Frost, C., et al. (2009). A meta-analysis of hippocampal atrophy rates in Alzheimer's disease. *Neurobiol. Aging* 30, 1711–1723. doi: 10.1016/j.neurobiolaging.2008.01.010

Beekly, D. L., Ramos, E. M., Lee, W. W., Deitrich, W. D., Jacka, M. E., Wu, J., et al. (2007). The National Alzheimer's coordinating center (NACC) database: the uniform data set. *Alzheimer Dis. Assoc. Disord.* 21, 249–258. doi: 10.1097/WAD.0b013e318142774e

Berezuk, C., Zakzanis, K. K., Ramirez, J., Ruocco, A. C., Edwards, J. D., Callahan, B. L., et al. (2017). Functional reserve: experience participating in instrumental activities of daily living is associated with gender and functional independence in mild cognitive impairment. *J. Alzheimers Dis.* 58, 425–434. doi: 10.3233/JAD-161227

Boccardi, M., Bocchetta, M., Ganzola, R., Robitaille, N., Redolfi, A., Duchesne, S., et al. (2015). Operationalizing protocol differences for EADC-ADNI manual hippocampal segmentation. *Alzheimers Dement.* 11, 184–194. doi: 10.1016/j.jalz.2013.03.001

Boutzoukas, E. M., O'Shea, A., Albizu, A., Evangelista, N. D., Hausman, H. K., Kraft, J. N., et al. (2021). Frontal white matter hyperintensities and executive functioning performance in older adults. *Front. Aging Neurosci.* 13:672535. doi: 10.3389/fnagi.2021.672535

Brickman, A. M., Schupf, N., Manly, J. J., Luchsinger, J. A., Andrews, H., Tang, M. X., et al. (2008). Brain morphology in older African Americans, Caribbean Hispanics, and whites from northern Manhattan. *Arch. Neurol.* 65, 1053–1061. doi: 10.1001/archneur.65.8.1053

Brickman, A. M., Zahodne, L. B., Guzman, V. A., Narkhede, A., Meier, I. B., Griffith, E. Y., et al. (2015). Reconsidering harbingers of dementia: progression of parietal lobe white matter hyperintensities predicts Alzheimer's disease incidence. *Neurobiol. Aging* 36, 27–32. doi: 10.1016/j.neurobiolaging.2014.07.019

Burton, R. L., O'Connell, M. E., and Morgan, D. G. (2018). Cognitive and neuropsychiatric correlates of functional impairment across the continuum of no cognitive impairment to dementia. *Arch. Clin. Neuropsychol.* 33, 795–807. doi: 10.1093/arclin/acx112

Burton, C. L., Strauss, E., Bunce, D., Hunter, M. A., and Hultsch, D. F. (2009). Functional abilities in older adults with mild cognitive impairment. *Gerontology* 55, 570–581. doi: 10.1159/000228918

Carmichael, O., Schwarz, C., Drucker, D., Fletcher, E., Harvey, D., Beckett, L., et al. (2010). Longitudinal changes in white matter disease and cognition in the first year of the Alzheimer disease neuroimaging initiative. *Arch. Neurol.* 67, 1370–1378. doi: 10.1001/archneurol.2010.284

DeCarli, C., Miller, B. L., Swan, G. E., Reed, T., Wolf, P. A., and Carmelli, D. (2001). Cerebrovascular and brain morphologic correlates of mild cognitive impairment in the National Heart, Lung, and Blood Institute twin study. *Arch. Neurol.* 58, 643–647. doi: 10.1001/archneur.58.4.643

DeCarli, C., Miller, B. L., Swan, G. E., Reed, T., Wolf, P. A., Garner, J., et al. (1999). Predictors of brain morphology for the men of the NHLBI twin study. *Stroke* 30, 529–536. doi: 10.1161/01.STR.30.3.529

DeCarli, C., Murphy, D. G., Teichberg, D., Campbell, G., and Sobering, G. S. (1996). Local histogram correction of MRI spatially dependent image pixel intensity nonuniformity. *J. Magn. Reson. Imaging* 6, 519–528. doi: 10.1002/jmri.1880060316

Ezzati, A., Katz, M. J., Lipton, M. L., Lipton, R. B., and Verghese, J. (2015). The association of brain structure with gait velocity in older adults: a quantitative volumetric analysis of brain MRI. *Neuroradiology* 57, 851–861. doi: 10.1007/s00234-015-1536-2

Farias, S. T., Lau, K., Harvey, D., Denny, K. G., Barba, C., and Mefford, A. N. (2017). Early functional limitations in cognitively normal older adults predict diagnostic conversion to mild cognitive impairment. *J. Am. Geriatr. Soc.* 65, 1152–1158. doi: 10.1111/jgs.14835

Fellows, R. P., and Schmitter-Edgecombe, M. (2018). Independent and differential effects of obesity and hypertension on cognitive and functional abilities. *Arch. Clin. Neuropsychol.* 33, 24–35. doi: 10.1093/arclin/acx045

Fellows, R. P., and Schmitter-Edgecombe, M. (2019). Multimethod assessment of everyday functioning and memory abilities in Parkinson's disease. *Neuropsychology* 33, 169–177. doi: 10.1037/neu0000505

Frisoni, G. Z., Bocchetta, M., Chételat, G., Rabinovici, G. D., De Leon, M. J., Kaye, J., et al. (2013). Imaging markers for Alzheimer disease: which vs how. *Neurology* 81, 487–500.

Gleason, C. E., Norton, D., Zuelsdorff, M., Benton, S. F., Wyman, M. F., Nystrom, N., et al. (2019). Association between enrollment factors and incident cognitive impairment in blacks and whites: data from the Alzheimer's disease center. *Alzheimers Dement.* 15, 1533–1545. doi: 10.1016/j.jalz.2019.07.015

Hu, H. Y., Ou, Y. N., Shen, X. N., Qu, Y., Ma, Y. H., Wang, Z. T., et al. (2021). White matter hyperintensities and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 36 prospective studies. *Neurosci. Biobehav. Rev.* 120, 16–27. doi: 10.1016/j.neubiorev.2020.11.007

Hwang, T. J., Masterman, D. L., Ortiz, F., Fairbanks, L. A., and Cummings, J. L. (2004). Mild cognitive impairment is associated with characteristic neuropsychiatric symptoms. *Alzheimer Dis. Assoc. Disord.* 18, 17–21. doi: 10.1097/00002093-200401000-00004

Kaufer, D. I., Cummings, J. L., Ketchel, P., Smith, V., MacMillan, A., Shelley, T., et al. (2000). Validation of the NPI-Q, a brief clinical form of the neuropsychiatric inventory. *J. Neuropsychiatry Clin. Neurosci.* 12, 233–239. doi: 10.1176/jnp.12.2.233

Kiselica, A. M., and Benge, J. F.Alzheimer's Disease Neuroimaging Initiative (2021). Neuropsychological equivalence of the clinical diagnosis of mild cognitive impairment in the National Alzheimer's coordinating center uniform data set and Alzheimer's Disease Neuroimaging Initiative. *Dement. Geriatr. Cogn. Disord.* 50, 231–236. doi: 10.1159/000516413

Kraal, A. Z., Massimo, L., Fletcher, E., Carrión, C. I., Medina, L. D., Mungas, D., et al. (2021). Functional reserve: the residual variance in instrumental activities of daily living not explained by brain structure, cognition, and demographics. *Neuropsychology* 35, 19–32. doi: 10.1037/neu0000705

McAlister, C., Schmitter-Edgecombe, M., and Lamb, R. (2016). Examination of variables that may affect the relationship between cognition and functional status in individuals with mild cognitive impairment: a meta-analysis. *Arch. Clin. Neuropsychol.* 31, 123–147. doi: 10.1093/arclin/acv089

McKay, N. S., Dincer, A., Mehrotra, V., Aschenbrenner, A. J., Balota, D., Hornbeck, R. C., et al. (2022). Beta-amyloid moderates the relationship between cortical thickness and attentional control in middle-and older-aged adults. *Neurobiol. Aging* 112, 181–190. doi: 10.1016/j.neurobiolaging.2021.12.012

Morris, J. C., Weintraub, S., Chui, H. C., Cummings, J., DeCarli, C., Ferris, S., et al. (2006). The uniform data set (UDS): clinical and cognitive variables and descriptive data from Alzheimer disease centers. *Alzheimer Dis. Assoc. Disord.* 20, 210–216. doi: 10.1097/01.wad.0000213865.09806.92

Overdorp, E. J., Kessels, R. P., Claassen, J. A., and Oosterman, J. M. (2016). The combined effect of neuropsychological and neuropathological deficits on instrumental activities of daily living in older adults: a systematic review. *Neuropsychol. Rev.* 26, 92–106. doi: 10.1007/s11065-015-9312-y

Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H. Jr., Chance, J. M., and Filos, S. (1982). Measurement of functional activities in older adults in the community. *J. Gerontol.* 37, 323–329. doi: 10.1093/geronj/37.3.323

Puzo, C., Labriola, C., Sugarman, M. A., Tripodis, Y., Martin, B., Palmisano, J. N., et al. (2019). Independent effects of white matter hyperintensities on cognitive, neuropsychiatric, and functional decline: a longitudinal investigation using the National Alzheimer's coordinating center uniform data set. *Alzheimers Res. Ther.* 11, 1–13. doi: 10.1186/s13195-019-0521-0

Rog, L. A., Park, L. Q., Harvey, D. J., Huang, C. J., Mackin, S., and Farias, S. T. (2014). The independent contributions of cognitive impairment and

neuropsychiatric symptoms to everyday function in older adults. *Clin. Neuropsychol.* 28, 215–236. doi: 10.1080/13854046.2013.876101

Rosano, C., Bennett, D. A., Newman, A. B., Venkatraman, V., Yaffe, K., Harris, T., et al. (2012). Patterns of focal gray matter atrophy are associated with bradykinesia and gait disturbances in older adults. *J. Gerontol. Ser. A: Biomed. Sci. Med. Sci.* 67, 957–962. doi: 10.1093/gerona/glr262

Sabuncu, M. R., Desikan, R. S., Sepulcre, J., Yeo, B. T. T., Liu, H., Schmansky, N. J., et al. (2011). The dynamics of cortical and hippocampal atrophy in Alzheimer disease. *Arch. Neurol.* 68, 1040–1048. doi: 10.1001/archneurol.2011.167

Salthouse, T. A. (2011). Neuroanatomical substrates of age-related cognitive decline. *Psychol. Bull.* 137, 753–784. doi: 10.1037/a0023262

Silbert, L. C., Dodge, H. H., Perkins, L. G., Sherbakov, L., Lahna, D., Erten-Lyons, D., et al. (2012). Trajectory of white matter hyperintensity burden preceding mild cognitive impairment. *Neurology* 79, 741–747. doi: 10.1212/WNL.0b013e3182661f2b

Silbert, L. C., Howieson, D. B., Dodge, H., and Kaye, J. A. (2009). Cognitive impairment risk: white matter hyperintensity progression matters. *Neurology* 73, 120–125. doi: 10.1212/WNL.0b013e3181ad53fd

Silbert, L. C., Nelson, C., Howieson, D. B., Moore, M. M., and Kaye, J. A. (2008). Impact of white matter hyperintensity volume progression on rate of cognitive and motor decline. *Neurology* 71, 108–113. doi: 10.1212/01.wnl.0000316799.86917.37

Teng, E., Becker, B. W., Woo, E., Cummings, J. L., and Lu, P. H. (2010). Subtle deficits in instrumental activities of daily living in subtypes of mild cognitive impairment. *Dement. Geriatr. Cogn. Disord.* 30, 189–197. doi: 10.1159/000313540

Vercauteren, T., Pennec, X., Perchant, A., and Ayache, N. (2007). Non-parametric diffeomorphic image registration with the demons algorithm. in *International conference on medical image computing and computer-assisted intervention* (Berlin, Heidelberg: Springer), 319–326.

Vermeulen, J., Neyens, J. C., van Rossum, E., Spreeuwenberg, M. D., and de Witte, L. P. (2011). Predicting ADL disability in community-dwelling elderly people using physical frailty indicators: a systematic review. *BMC Geriatr.* 11, 1–11. doi: 10.1186/1471-2318-11-33

Wang, H., Suh, J. W., Das, S. R., Pluta, J. B., Craige, C., and Yushkevich, P. A. (2012). Multi-atlas segmentation with joint label fusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 611–623. doi: 10.1109/TPAMI.2012.143

Weintraub, S., Salmon, D., Mercaldo, N., Ferris, S., Graff-Radford, N. R., Chui, H., et al. (2009). The Alzheimer's disease centers' uniform data set (UDS): the neuropsychological test battery. *Alzheimer Dis. Assoc. Disord.* 23, 91–101. doi: 10.1097/WAD.0b013e318191c7dd

Willey, J. Z., Scarmeas, N., Provenzano, F. A., Luchsinger, J. A., Mayeux, R., and Brickman, A. M. (2013). White matter hyperintensity volume and impaired mobility among older adults. *J. Neurol.* 260, 884–890. doi: 10.1007/s00415-012-6731-z

Zahodne, L. B., and Tremont, G. (2013). Unique effects of apathy and depression signs on cognition and function in amnestic mild cognitive impairment. *Int. J. Geriatr. Psychiatry* 28, 50–56. doi: 10.1002/gps.3789

Zhao, L., Matloff, W., Ning, K., Kim, H., Dinov, I. D., and Toga, A. W. (2019). Age-related differences in brain morphology and the modifiers in middle-aged and older adults. *Cereb. Cortex* 29, 4169–4193. doi: 10.1093/cercor/bhy300

10.3389/fnagi.2022.952145

Appendix

Partial correlations between pathological functioning and cognitive variables in MCI group.

	$r_{partial}$	p
LM I – Immediate Recall	−0.178	0.023
LM II – Delayed Recall	−0.261	<0.001
Digit Span Forward	−0.042	0.597
Digit Span Backward	0.052	0.509
Animal Fluency	−0.164	0.036
Vegetable Fluency	−0.243	0.002
Digit Symbol	0.047	0.548
Boston Naming Test	−0.257	<0.001
Trails A	0.133	0.090
Trails B	−0.034	0.664

Partial correlations are adjusted for age, sex, race/ethnicity, education, and APOE ε4 carrier status.

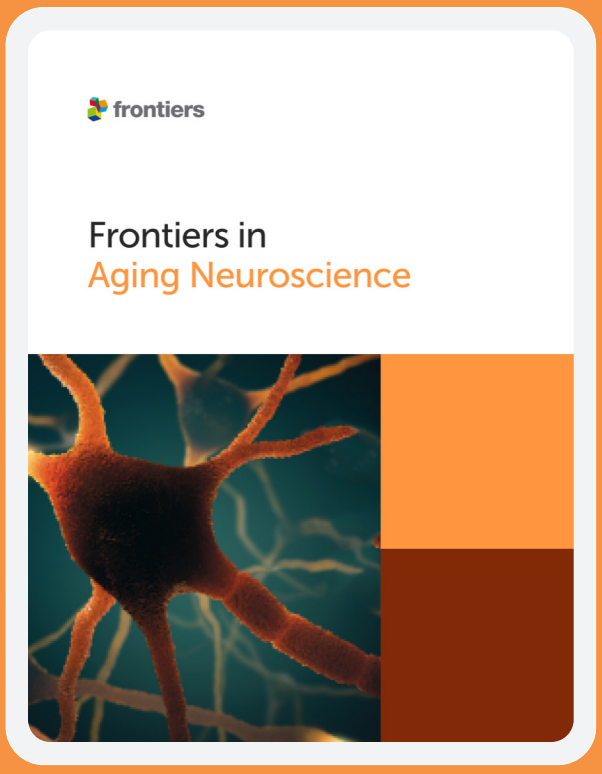
Frontiers in
Aging Neuroscience

Explores the mechanisms of central nervous system aging and age-related neural disease

The third most-cited journal in the field of geriatrics and gerontology, with a focus on understanding the mechanistic processes associated with central nervous system aging.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact