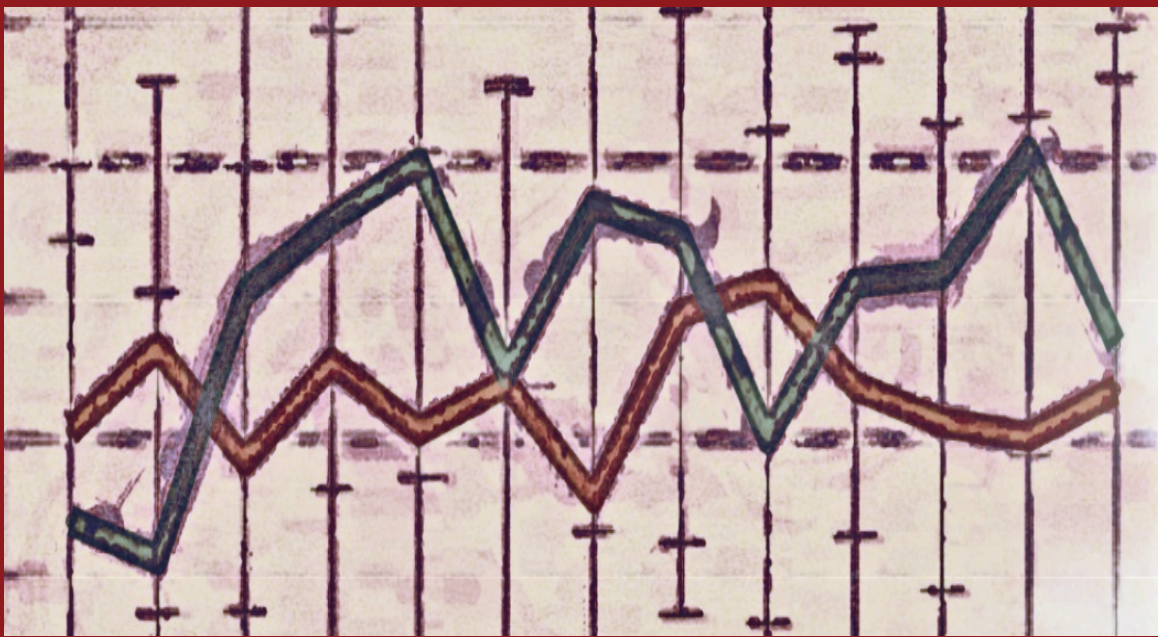
UNIVERSITY OF COPENHAGEN
FACULTY OF HEALTH AND MEDICAL SCIENCES

How to use physical activity monitoring to enhance physical activity levels among older adults



PhD Thesis

Rasmus Tolstrup Larsen, PT, MSc

University of Copenhagen
Faculty of Health and Medical Sciences
Department of Public Health
Section of Social Medicine

This thesis has been submitted to the Graduate School of Health and Medical Sciences, University of Copenhagen, December 23, 2020.

Name of institute: Department of Public Health

Name of department: Section of Social Medicine

Author: Rasmus Tolstrup Larsen

Title: How to use physical activity monitoring to enhance physical activity levels among older adults.

Dansk titel: Hvordan fysisk aktivitetsmonitorering kan bruges til at øge det fysiske aktivitetsniveau hos ældre.

Subject description: Physical activity monitoring holds the potential of serving as facilitators of healthy behavioural change through feedback and goal setting among individuals and the ability to serve as outcome instruments in clinical settings. This thesis provides evidence for how physical activity monitors can be used among older adults.

Supervisors: Primary supervisor: Professor Henning Langberg
Primary co-supervisor: Professor Emeritus Henning Boje Andersen
Co-supervisors: Associate professor Carsten Bogh Juhl and Senior Researcher Jan Christensen

Submission date: December 23, 2020

Word count: 22,562

Correspondence

Rasmus Tolstrup Larsen

University of Copenhagen, Faculty of Health and Medical Sciences, Department of Public Health, Section of Social Medicine, Gothersgade 160, 3rd floor, 1123, Copenhagen K, Denmark

E-mail: rala@sund.ku.dk

Academic advisors

1) Professor Henning Langberg, Section of Health Services Research, Department of Public Health, University of Copenhagen, Denmark.
2) Professor Emeritus Henning Boje Andersen, Technical University of Denmark, Department of Technology, Management and Economics.
3) Associate Professor Carsten Bogh Juhl, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark.
4) Senior Researcher Jan Christensen, Department of Occupational Therapy and Physiotherapy, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark.

Assessment committee

Chair:

Associate Professor Signe Smith Jervelund, Danish Research Centre for Migration, Ethnicity and Health (MESU), Department of Public Health, University of Copenhagen, Copenhagen, Denmark.

Opponents:

Professor Ulf Ekelund, Department of Sport Medicine, Norwegian School of Sport Sciences, Oslo, Norway.

Professor Ulrik Dalgas, Section of Sport Science, Department of Public Health, Aarhus University, Aarhus, Denmark.

The Oral Defence takes place on April 9, 2021.

TABLE OF CONTENTS

PREFACE AND ACKNOWLEDGEMENTS

The research presented in this PhD thesis was carried out at the section of Social Medicine at the Department of Public Health, University of Copenhagen. The research was funded by the EU Horizon 2020 program, the REACH2020 consortium and Copenhagen Centre for Health Technology.

First of all, I would like to thank my best friend, life partner and the mother to my daughter, Stephanie, for inspiring me to pursue a career in research and academia. Without her, I would not have started my master's program, not to mention my PhD. Furthermore, having the opportunity to have a fellow PhD student at home has given me the opportunity to share ups and downs with a peer. Especially, her ability to tell me when to stop starting new projects is probably the main reason this thesis was handed in by the deadline.

Off course, I would like to thank my primary supervisor, Professor Henning Langberg. With Henning Langberg's style of leadership and ability to begin new projects, I have always felt at home, at ease, and able to work effectively. No matter if it was at late research meetings, Wiener schnitzel and Weissbier in Munich or "data collection" at Roskilde Festival, it has always been a fun and productive ride. I would like to thank Professor Henning Boje Andersen for always providing me with interesting new insights in the field of ageing research and technology, but also for practicing my German with me and recommending great novels. I would like to thank Associate Professor Carsten Bogh Juhl for doing exactly what he was supposed to do in this PhD project. Carsten is the best research mentor one can get and has always thoroughly and rigorously supervised my research and statistical methods. Lastly, from my group of supervisors, I should also mention Jan Christensen. Jan has spent more hours on me than he was allocated from the PhD school. Jan is always prompt to respond, very thorough and not afraid of taking part of "the manual work." It has been a pleasure to work as his research assistant, fellow PhD student and since to have him supervise my project. A lot of thinking and discussing have been done at lunches or over a beer.

I would also like to thank my primary research assistant and friend, Christoffer Bruun Korfitsen for being a priceless aid, and always reliable and steady going fellow researcher. Christoffer has been part of most of my PhD studies, especially a very big part of the RCT, but he has also been part of almost every other project we have started during my short research career.

Without him, this PhD project would have looked entirely different, and would definitely not have been as comprehensive.

I would also like to thank Professor John Hirdes and Postdoctoral Researcher Luke Turcotte for welcoming me at the department of Health Services Research at University of Waterloo, Ontario Canada. During my research stay, I learned to use new software, new registers and to collaborate with foreign colleagues. I really appreciate our work together.

Finally, I would like to thank John, Paul, George and Ringo for motivating me in my contemplation, preparation, and action phase, as well as for inspiring and educating those who helped me in my maintenance phase avoiding termination (no one mentioned, no one forgotten).

LIST OF STUDIES

This PhD thesis consists of the following synopsis and includes six individual papers listed below:

Study I:

Paper A:
Larsen, Rasmus Tolstrup, Jan Christensen, Carsten Bogh Juhl, Henning Boje Andersen, and Henning Langberg. "Physical Activity Monitors to Enhance the Daily Amount of Physical Activity in Elderly—a Protocol for a Systematic Review and Meta-Analysis." *Systematic Reviews* 7, no. 1 (May 2, 2018): 69. https://doi.org/10.1186/s13643-018-0733-6.

Paper B:
Larsen, Rasmus Tolstrup, Jan Christensen, Carsten Bogh Juhl, Henning Boje Andersen, and Henning Langberg. "Physical Activity Monitors to Enhance Amount of Physical Activity in Older Adults – a Systematic Review and Meta-Analysis." *European Review of Aging and Physical Activity* 16, no. 1 (May 4, 2019): 7. https://doi.org/10.1186/s11556-019-0213-6.

Study II:

Paper C:
Larsen, Rasmus Tolstrup, Christoffer Bruun Korfitsen, Carsten Bogh Juhl, Henning Boje Andersen, Henning Langberg, and Jan Christensen. "Criterion Validity for Step Counting in Four Consumer-Grade Physical Activity Monitors among Older Adults with and without Rollators." *European Review of Aging and Physical Activity* 17, no. 1 (December 2020): 1. https://doi.org/10.1186/s11556-019-0235-0.

Study III:

Paper D:
Larsen, Rasmus T., Korfitsen, Christoffer B., Juhl, Carsten B., Boje Andersen, Henning, Langberg, Henning, & Christensen. Jan. (2020). Concurrent Validity Between Electronically Administered Physical Activity Questionnaires and Objectively Measured Physical Activity in Danish Community-Dwelling Older Adults. *Journal of Aging and Physical Activity*. Advance online publication. https://doi.org/10.1123/japa.2020-0214.

Study IV:

Paper E:
Larsen RT, Korfitsen CB, Juhl CB, Andersen HB, Christensen J, Langberg H. The MIPAM trial: a 12-week intervention with motivational interviewing and physical activity monitoring to enhance the daily amount of physical activity in community-dwelling older adults – a study protocol for a randomized controlled trial. BMC Geriatrics. 2020;20(1). 412. https://doi.org/10.1186/s12877-020-01815-1.

Paper F (under review in European Journal of Aging and Physical Activity):
Larsen, Rasmus Tolstrup, Christoffer Bruun Korfitsen, Camilla Keller, Carsten Bogh Juhl, Henning Boje Andersen, Henning Langberg, and Jan Christensen. "The MIPAM trial: A 12-week intervention with motivational interviewing and physical activity monitoring, to enhance the daily amount of physical activity in community-dwelling older adults.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

(**95%CI**) 95% Confidence Interval

(**BMI**) Body Mass Index

(**COSMIN**) Consensus-Based Standards for selection of Health Measurement Instruments

(**CONSORT**) Consolidated Standards of Reporting Trials

(**COMET**) Core Outcome Measures in Effectiveness Trials

(**EU**) European Union

(**EQ-5D-5L**) EuroQol-5 Domain

(**HRQoL**) Health Related Quality of Life

(**ICC**) Interclass Correlation Coefficient

(**IPAQ-SF**) International Physical Activity Questionnaire Short Form

(**IQR**) Interquartile Range

(**MET**) Metabolic Equivalent of Task

(**MVPA**) Moderate to Vigorous Physical Activity

(**MI**) Motivational Interviewing

(**NPAQ-short**) Nordic Physical Activity Questionnaire Short

(**OEE**) Outcome-Expectancy for Exercise

(**PA**) Physical Activity

(**PAM**) Physical Activity Monitors

(**PAQ**) Physical Activity Questionnaire

(**PRISMA**) Preferred Reporting Items for Systematic Reviews and Meta-Analyses

(**QoL**) Quality of Life

(**SEE**) Self-Efficacy for Exercise

(**SD**) Standard Deviation

(**SMD**) Standardised Mean Differences

(**TIDier**) Template for Intervention Description and Replication

(**PROSPERO**) The International Prospective Register of Systematic Reviews

(**UCLA**) University of California, Los Angeles

(**WHO**) World Health Organization

ENGLISH SUMMARY

Inadequate physical activity (PA) levels have a major impact on public health, and inactivity constitutes a key barrier to healthy aging among older adults. The least physically active older adults may benefit from even small increases in PA levels, so effective interventions should be prioritised. Innovative solutions should be implemented for PA-promoting interventions aimed at older adults to ensure that they are relevant, feasible, and cost-effective. PA Monitors (PAMs) are in common use to facilitate behavioural change, but the literature is highly variable as to their effectiveness among older adults. As one size does not fit all, this thesis suggests further behavioural change methods in combination with PAM-based interventions.

Motivational Interviewing (MI) is a well-established method in Denmark and has been reported to increase PA levels among older adults, and thus it is relevant as an additional intervention when used in combination with PAMs. Most consumer-available PAMs have the potential to be used as intervention content as well as outcome instruments. The validity and measurement properties of research-grade PAMs are already well established, but before consumer-available PAMs are used in clinical research, the measurement properties of each individual model in each specific population should be investigated. PA Questionnaires (PAQs) are frequently used to aid or replace PAMs in clinical studies, but though this is a feasible approach in most situations, their measurement properties can pose a challenge. Hence, the measurement properties and, especially, the validity of the PAQs among older adults should be investigated before they are decided upon for a research project. To investigate how PA monitoring can be used to enhance PA levels among older adults, four studies were conducted. A systematic review and meta-analysis investigated the effectiveness of PAMs among older adults, while two validity studies investigated the measurement properties of consumer-available PAMs and two PAQs. Finally, a randomised controlled trial (RCT) investigated the effectiveness of MI as an add-on intervention to a PAM-based intervention among older adults.

In summary, the systematic review demonstrated low quality of evidence for a clinically highly relevant moderate effect of PAM-based intervention. The criterion validity of consumer-grade PAMs was found to be mainly affected by position of use, as hip-worn PAMs did not differ in terms of measurement error or validity, whereas wrist-worn monitors were found to be associated with inadequate validity and a high degree of measurement error. Among older adults, the concurrent validity of two PAQs was found to be inadequate and their use could not be

supported. When added to a PAM-based intervention, MI was found to be effective, but non-significant, in increasing the daily step count among older adults.

The thesis includes six peer-reviewed publications from four studies to answer PA monitoring among older adults can be used. The findings should be interpreted within the limitations from the four studies, including heterogeneity and low quality of evidence in the systematic review, low statistical power from lack of participants in the RCT, and several types of bias affecting PAMs in general. However, in conclusion, PAMs seem effective, feasible and safe to use in the population of older adults, and the thesis points to several implications for further research and clinical practice.

DANSK RESUMÉ

Fysiske aktivitetsniveau har stor betydning for folkesundheden. Inaktivitet er en af de primære barrierer for sund aldring, og især blandt de mindst aktive har små forskelle i det fysiske aktivitetsniveau en stor betydning for sundheden. Derfor bør effektive interventioner prioriteres. Interventionerne skal ikke kun være effektive, men også innovative, da anvendeligheden og effektiviteten i forhold til omkostningerne altid skal balanceres, før interventionerne kan implementeres. Fysisk aktivitetsmonitorering bliver hyppigt anvendt som et redskab i forbindelse med adfærdsændringer, men især blandt ældre er evidensen ikke entydig. Ikke alle ældre reagerer på samme måde, når de modtager feedback fra aktivitetsmonitorer, og derfor foreslås det at yderligere adfærdsteorier implementeres og afprøves i forbindelse med fysisk aktivitetsmonitorering. Den motiverende samtale er en samtaleteknik, der målrettes adfærdsændringer hos den enkelte, og er implementeret i flere sektorer i Danmark. Den motiverende samtale har yderligere vist sig at være et effektivt redskab i forhold til fysisk aktivitet blandt ældre. Derfor bør effekten af den motiverende samtale undersøges i forbindelse med en intervention, der består af fysisk aktivitetsmonitorering blandt ældre.

Aktivitetsmonitorer, der er designet og godkendt til klinisk forskning, har alle veletablerede måleegenskaber, mens egenskaberne for aktivitetsmonitorer, der er tilgængelige for forbrugere, ikke har samme udgangspunkt. De aktivitetsmonitorer, der er tilgængelige for forbrugere, har dog den fordel, at de kan samtidig bruges som et redskab til adfærdsændringer på grund af deres feedback til brugeren og som et måleinstrument. Men før disse kan anvendes i klinisk forskning, bør måleegenskaberne dog undersøges nærmere. For at bistå aktivitetsmonitorer er spørgeskemaer, der klarlægger fysisk aktivitet, blevet anvendt længe. Selvom de er en billig og nem løsning, er spørgeskemaerne hæmmet af forskellige systematiske målefejl. Derfor bør måleegenskaberne hos disse spørgeskemaer blandt ældre også undersøges nærmere.

For at undersøge hvordan fysisk aktivitetsmonitorering kan anvendes til at øge det fysiske aktivitetsniveau blandt ældre blev fire studier udført i forbindelse med dette ph.d.-projekt. En systematisk oversigtsartikel og meta-analyse blev gennemført for at undersøge den samlede effekt af feedback fra fysisk aktivitetsmonitorering målt på ældres fysiske aktivitetsniveau. To valideringsstudier blev gennemført: Ét for at undersøge måleegenskaberne hos aktivitetsmonitorer, der er tilgængelige for forbrugere, og ét for at undersøge måleegenskaberne hos spørgeskemaer, der klarlægger fysisk aktivitet blandt ældre. Slutteligt blev et randomiseret kontrolleret forsøg

gennemført for at undersøge effekten af den motiverende samtale blandt ældre, når den kombineres med en intervention, der anvender fysisk aktivitetsmonitorering.

Lav kvalitet af evidens blev fundet for en klinisk relevant moderat effekt af fysisk aktivitetsmonitorering blandt ældre. Kriterievaliditeten af fire aktivitetsmonitorer, der er tilgængelige for forbrugere, blev primært påvirket af placeringen. De håndledsbårne monitorer havde en stor grad af måleusikkerhed og havde derfor ikke de fornødne måleegenskaber blandt en population af ældre, der både inkluderer rollatorbrugere, og ældre, der går uden hjælpemidler. Den samstemmende validitet af to spørgeskemaer, der klarlægger fysisk aktivitet, blev fundet utilstrækkelig, og spørgeskemaerne bør derfor ikke anvendes blandt ældre uden grundig gennemgang af indholdet med den enkelte respondent. Den motiverende samtale blev fundet effektiv, dog ikke signifikant, i at øge ældres fysiske aktivitetsniveau, når den blev kombineret med fysisk aktivitetsmonitorering.

Denne ph.d.-afhandling inkluderer seks fagfællebedømte artikler fra fire studier for at svare på, hvordan fysisk aktivitetsmonitorering kan anvendes til at øge ældres fysiske aktivitetsniveau. Afhandlingens fund bør tolkes i lyset af de begrænsninger, der diskuteres, herunder lav kvalitet af evidens og variation i den systematiske oversigtsartikel, lav statistisk præcision i det randomiserede kontrollerede forsøg og flere typer af bias, der kan påvirke fysisk aktivitetsmonitorering generelt. Slutteligt ses fysisk aktivitetsmonitorering dog som relevant, effektivt og anvendeligt blandt ældre, hvorfor denne afhandling opstiller flere implikationer for fremtidig forskning og klinisk praksis.

THESIS AT A GLANCE

Study # and title	Objectives	Design and Methods	Results and conclusion
Study I (paper A and B) **Physical Activity Monitors to enhance amount of physical activity in older adults – a systematic review and meta-analysis.**	To estimate the effect of PAM-based interventions on PA behaviour in participants aged 65 and above.	**Design: systematic review and meta-analysis** Searches in five databases were performed. RCTs and randomised cross-over trials were included. Random-effects meta-analysis using Hedges' g, were used to pool the study results.	Twenty-one studies with 2,783 participants were included. With low quality of evidence, PAM-based interventions had a moderate effect (SMD=0.54, 95% CI: 0.34 to 0.73) compared to control interventions, corresponding to an average increase of 1,297 steps per day in the intervention groups. No impact of patient and intervention characteristics on the effect estimates were found.
Study II (paper C) **Criterion validity for step counting in four consumer-grade Physical activity monitors among older adults with and without rollators.**	To investigate the criterion validity of four consumer-grade PAMs in older adults and to investigate whether the measurement properties were affected by placement and assistive devices.	**Design: validity study** Participants performed self-paced walking while their steps were visually counted. The participants wore 16 monitors (four from each device; Misfit Shine, Nokia GO, Jawbone UP Move and Garmin Vivofit 3). ICC (1,2) were used to express criterion validity.	A total of 103 older adults participated. The Nokia GO was excluded due to technical issues. The hip-worn PAMs did not differ significantly in terms of measurement error or criterion validity. Wrist-worn PAMs cannot adequately measure number of steps in a population of older adults using rollators. The hip-worn PAMs were superior to wrist-worn PAMs among older adults with and without rollators.
Study III (paper D) **Criterion Validity between Electronically Administered Physical Activity Questionnaires and Objectively Measured Physical Activity in Danish Community-Dwelling Older Adults.**	To investigate the concurrent validity between the IPAQ-SF and the NPAQ-Short and objectively measured daily steps among older adults.	**Design: validity study** Using baseline data from participants enrolled in study IV, the MIPAM trial, Spearman's rho was used to express concurrent validity.	Fifty-four participants were included. In general, both questionnaires performed inadequately with only two sub-scales showing moderate to good correlation with daily steps. The concurrent validity was low as the scores did not reflect objectively measured daily steps.
Study IV (paper E and F) **The MIPAM trial – Motivational Interviewing and Physical Activity Monitoring to enhance the daily Level of Physical Activity among Older Adults – a Randomised Controlled Trial.**	To investigate the effect of MI as an add-on intervention to a PAM-based intervention measured in community-dwelling older adults.	**Design: RCT** Two-arm parallel group RCT investigating if a PAM+MI intervention were more effective than a PAM intervention. Average daily step count, was analysed following the ITT principle with multiple imputations.	In total, 70 participants were included. During the intervention period, the PAM+MI group walked on average 909 more steps per day than PAM group, however insignificant (95%CI: -71; 1889) and reported 2.3 points less on the UCLA Loneliness Scale (95%CI: -4.5; -1.24).

Abbreviations: Physical Activity Monitor (PAM), Randomised Controlled Trial (RCT), Standardised Mean Difference (SMD), Interclass Correlation Coefficient (ICC), International Physical Activity Questionnaire Short Form (IPAQ-SF), Nordic Physical Activity Questionnaire Short (NPAQ-Short), Motivational Interviewing and Physical Activity Monitoring (MIPAM), Motivational Interviewing (MI), Intention-To-Treat (ITT), 95% Confidence Interval (95%CI), University of California, Los Angeles (UCLA).

INTRODUCTION

For the first time in history, the majority of all living people can expect to live into their sixties (1). A longer life equals more opportunities for older adults and their families, with the possibility of pursuing new activities, further education and new ways of participating in society. However, this is dependent on the health the individual (1). The World Health Organization (WHO) defines healthy aging as "*the process of developing and maintaining the functional ability that enables wellbeing in older age*" (2). Chances of healthy or successful aging depend on several individual risk factors, as well as social functioning, and are thus dependent on health and social systems and levels of welfare provision (2–4).

Changing demographics and the physical activity pandemic

According to the European Commission's 2018 Aging Report, population ageing and demographic changes will turn the European Union (EU) "*increasingly grey in the coming decades*." The total population in the EU is expected to increase by 9 million from 2016 to 2070, while the working age population will decrease with 41 million due to lower fertility, increasing life expectancy and migration flows (5). According to data from the World Bank, the life expectancy of a new-born has increased globally from 67.5 years in 2000 to 72.6 in 2018 and from 76.6 years to 81.4 in Denmark (6). Furthermore, the percentage of the population aged 65 and above has increased from 5.0% to 8.9% globally and from 10.6% to 19.8% in Denmark (7). All countries face major structural challenges as the pace of population aging is faster than ever before. Health and social systems should be ready for this demographic shift (1).

Noncommunicable diseases account for 71% of all deaths globally and are consequently also the major burden of most health systems (8,9). Physical inactivity has a major impact on global public health, as it is responsible for 9% of premature death globally (10). Among older adults, physical inactivity is associated with a trajectory towards disability, disease and premature death, and thus constitutes a barrier to healthy aging (3,11). Physical inactivity is not only a key risk factor for the individual older adult, but its high prevalence is also a problem for European societies across the board, as the prevalence of physical inactivity among older adults ranges from 4.9% in Sweden to 29% in Portugal (10,12). In Denmark, one in four people above the age of 65 and one in two above the age of 75 do not meet the WHO recommendations for minimum physical activity (PA) (13,14). It is estimated that physical inactivity is accountable for a substantial global economic burden of 47.8 billion euro, most of which borne by the public sector (15,16). It is well established that PA levels are highly associated with all-cause mortality, and recent evidence shows that it has a non-linear relationship with health, as even small changes in PA levels may benefit the least active older adults who stand to benefit the most from increasing their PA levels (17,18). Exercise among physically inactive older adults can be used to improve physical capacity, daily functioning and independence and hence as adjunct treatment for preventive measures and rehabilitation (11,19,20). Furthermore, older adults, including those with chronic diseases, will gain longevity benefits by engaging in more PA, no matter how physically active they used to be, which increases their chances of healthy aging (3,21). In summary, inactivity among older adults is a global burden. However, it is a problem that also has potential solutions on the individual level as well as from a public health perspective, because even small changes will have an impact on the risk of all-cause mortality (10,17,22).

Motivation and innovative physical activity programs

PA levels are inversely associated with age (21), which is expected due to the functional decline associated with ageing. However, despite the lower levels of PA among older adults compared to younger populations, older adults are often quite motivated to engage in PA or exercise and the vast majority of their reported favourite leisure time activities are somewhat active (23). PA-promoting interventions among older adults hold some potential, as a systematic review and meta-analysis reported moderate evidence for small to moderate effect sizes favouring intervention groups (24,25).

To effectively reverse or stop the global burden of inactivity, PA-promoting public health interventions should be upscaled to increase their ability to enhance levels of PA in different populations across varying cultural, geographical, social and economic contexts and to align with the United Nations Sustainable Development Goals (26). The WHO publication "Global Action Plan on Physical Activity 2018-2030" also addresses this topic and aims to reduce the global prevalence of physical inactivity by creating active societies, environments, people and systems (27). Ambulant and transportation activities should play a key role and serve as one of the main constructs in the interventions (28,29). Walking in particular is described as the most common type of PA among older adults, the vast majority of whom use walking both for transportation and exercise purposes (30,31). Furthermore, the amount of walking and number of daily steps play a key role in preventive medicine as they are strongly inversely associated with premature all-cause mortality (17,22), which makes walking and daily steps a key construct to be targeted in PA-promoting interventions among older adults. For PA-promoting interventions targeting older adults to be applicable, feasible and cost-effective, solutions that are innovative both in design and conceptualisation need to be implemented.

The Internet of Things is a network that enables connectivity and collection and exchange of data between electronic devices (32). This category of devices includes PA Monitors (PAMs), which refers to all types of devices that measure PA, such as pedometers, accelerometers and heart rate monitors (32,33). Quantifying PA with PAMs for research purposes has been done since the 1960s (34). But on top of being able to track and measure PA, modern PAMs are also being used to facilitate behavioural change, as they can provide users with feedback on their amount of PA (35). As early as 2007, a systematic review reported that PAMs could be effectively used to motivate and facilitate increased PA levels among adults (36). However, older adults might react differently to feedback on activity levels, and randomised controlled trials have reported PAM-based PA interventions to be effective (37–42) while others have reported them to be ineffective (43–47). The above-mentioned studies exhibit a large clinical, methodological- and statistical heterogeneity, so the literature on the effectiveness of PAM-based PA promoting interventions targeted populations of older adults need to be systematically reviewed.

If PAMs are shown to be an effective way to facilitate PA behavioural change within older adults, future studies should investigate whether add-on interventions can increase the effect of PAMS and PA adherence among older adults. Interventions based on behavioural change theories have the potential to change behaviour effectively (48,49). Goal setting, self-monitoring, action planning, and education on behaviour-health link models have been reported to be important in PA-interventions (49–53). Motivational Interviewing (MI) uses empathic listening, encouraged self-reflection and active counselling to guide individuals (54). MI aims to encourage behavioural change through motivation and self-efficacy (55,56). MI is already well-established and commonly used in many health- and social interventions in Denmark (57–63). MI have been reported to increase PA levels among older adults with hearth failure (64) and hip fracture (65). Older adults are already reported find the combination of PAM-based PA interventions and MI acceptable (66), so this specific behavioural change theory would appear highly relevant to use as an add-on intervention to PAMs within the population of Danish older adults.

Physical activity monitors as intervention facilitators and measurement instruments

As previously mentioned, PAMs have a promising role as facilitators in individual behavioural change interventions, and thus also in public health behavioural change interventions (35,67). Furthermore, modern PAMs are able to both track PA levels and collect longitudinal data, which enables the user not only to have real time or daily feedback on PA levels, but also to show trajectories and averages over longer period of time (35,68). The requirements for validity and measurement properties of the PAMs are not particularly stringent if they are only used to facilitate behavioural change, but if the PAMs are to be used in clinical research to measure outcomes, adequate measurement properties are very important. The validity and measurement properties of research grade-accelerometers have already been established as high (69). However, most research-grade accelerometers are sealed and do not provide the user with feedback on PA, and hence they are not ideal for facilitating behavioural change among individuals. Consumer-available PAMs (Fitbits, Apple Watches, Garmin Fitness Trackers, etc.) are used more frequently than ever among the general population (35,67) and have the potential to serve as both the intervention content as well as outcome instruments in intervention studies. However, before using consumer-available PAMs are put to use in clinical research in this dual role as behavioural change facilitators and outcome instruments, the measurement properties and, in particular, the criterion validity of specific PAMs should be evaluated. Measurement properties are expected to differ between populations, but especially so between older adults and younger populations as many older adults have a different gait cycle (70). As the algorithms for most consumer-available PAMs have been developed in healthy younger populations (33) and factors such as stride length, gait speed and use of assistive devices have been found to affect the validity of PAMs (70,71), the measurement properties of consumer-available PAMs should be investigated in the target population of older adults before they are used in studies. As only few studies have investigated their measurement properties among older adults (71–78), and none of these have studied whether the results of specific PAMs differed with placement (hip or wrist), it would appear highly relevant to include these specific questions in future research.

Within the field of PA measurement, there is something of an inverse association between precision and validity and the feasibility and practicality of the instruments. To accurately measure total energy expenditure, doubly labelled water is considered the gold standard, but it is rarely feasible to use at scale (79). For this reason, accelerometery and pedometers are often used instead to provide objective measurements of PA (80,81). Despite being accurate and valid for measuring PA and steps per day respectively, accelerometers and pedometers will only measure PA when worn, and studies are encouraged to include wear time in their analyses (82). Because of their time and cost efficiency, participant-reported measures of PA, typically in the form of PA recall questionnaires, are often used as an alternative to objective measurement (83). Although being feasible in most studies and situations, the psychometric properties, and thus the measurement properties, are challenged by recall and social desirability biases (84–87). Among the many PA questionnaires (PAQs) developed through the last decades are the commonly used International Physical Activity Questionnaire Short Form Version (IPAQ-SF) (88,89), and the Danish alternative, the Nordic Physical Activity Questionnaire (NPAQ-short) (14,90). When conducting a PA-promoting intervention study, the above-mentioned PAQs would appear relevant to include to investigate any between group differences not captured by the PAMs. Still, the measurement properties and – especially – the validity of the IPAQ-SF and the NPAQ-Short among older adults remain somewhat unclear (86,91–94).

Aims and objectives

This PhD project has investigated how PAMs can be used to enhance PA levels among older adults. Consequently, this project aimed to answer the following research questions:

1) How effective are PAM-based interventions on PA behaviour among older adults?
2) What is the criterion validity of consumer-grade PAMs among older adults with and without assistive devices?
3) What is the concurrent validity of PAQs among older adults?
4) How effective is MI as an add-on intervention to a PAM-based intervention in community-dwelling older adults?

BACKGROUND

This section presents the theoretical framework for the thesis and includes population and healthy ageing, ageing processes and functional decline, global PA recommendations, PA monitoring, behavioural change, and evidence-based research.

Population and healthy ageing

Population ageing refers to overall the ageing of the world's populations, a phenomenon that affects almost all countries, whereby both the total number and the proportion of older people are increasing (1,95). As previously described, here illustrated by the European Commission Statistical Office in Figure 1 below, many developed countries are approaching the first time in history where they will have a larger proportion of older adults than younger adults (1,95). The old-age dependency ratio, defined as the number of persons aged 65 or above divided by the number of adults aged 15 to 64, is one of the most used indicators for changing demographics. It is rapidly increasing in almost all regions of the world (95,96). This demographic shift will have a considerable societal impact across the economy, healthcare, welfare, social systems, labour markets, family structures and many other areas (95). However, even though the rapidly increasing old-age dependency ratio will pose a considerable financial challenge, as older adults are expected to live longer, older adults will also have the opportunity to continue to engage in societal activities and employment (1,95,97).

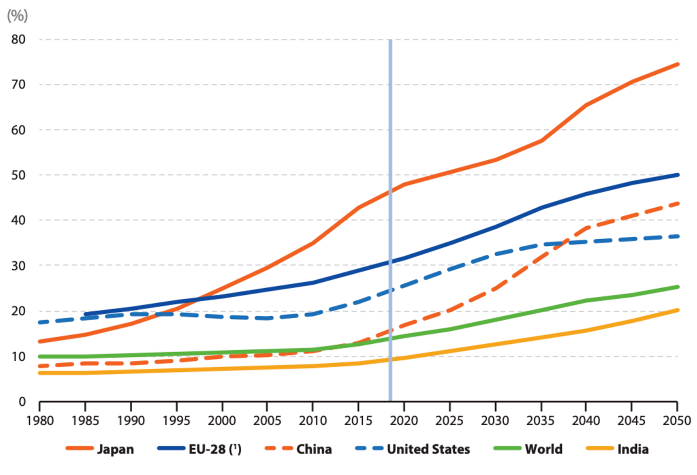


Figure 1. Old-age dependency ratio, 1980-2050 from the European Commission Statistical Office Publication: Ageing Europe 2019 Edition (95). The old-age dependency ratio is calculated as the number of people aged ≥ 65 years divided by the number of people aged 15-64 years, expressed as a percentage.

As early as 1982, the first World Assembly on Ageing addressed this issue by calling for action within many areas, such as research, health and nutrition, housing and environment, social welfare, income security and employment (98). These societal action plans were complex and not all directly relevant to the aim of the thesis. However, the foundation of all action plans was to engage older adults in societal activities and create better opportunities for healthy ageing (98). To this date, healthy ageing and successful ageing is still targeted at both public and individual levels (2,97). Healthy and successful ageing are two sides of the same coin. Both constructs are multi-dimensional where successful ageing is defined as being free from disease, physical and cognitive disability, which distinguishes it from the expected age-related trajectory towards decline in physical and cognitive function (99). Healthy ageing, as previously defined, is different from successful ageing in that it is more inclusive and does not require disease avoidance (2,99). There are four major requirements for healthy ageing; a change in the way we think about ageing in general, change in environments, alignment of health systems to the needs of older adults and development of systems for long-term care (1,2). Because of this, it is not only up to the individual older adult to stay healthy when old age approaches, but also a societal responsibility to facilitate healthy aging among individuals by adhering to the requirements and creating the best possibilities for older adults to be able to do the things they value for as long as possible (1,2).

Ageing processes and functional decline

There are several theories that seek to explain the pathological or clinical findings found with the inevitable deterioration associated with ageing on different levels (100,101). These include but are not limited to molecular and biological theories of ageing (102–104), psychological theories of ageing (105), and social theories of ageing (106). No matter the theory, on what level it operates or how it seeks to explain the process of ageing, the functional and cognitive decline associated with ageing is real, and the increasing risk of morbidity and frailty is associated with the level of PA (107–109). A systematic review by Paterson et al. concluded from 66 independent prospective cohort studies with functional or cognitive outcomes that regular aerobic activity and short-term exercise interventions are associated with a reduced risk of functional limitations and disability in older age (110). The findings are illustrated in Figure 2 below, where the level of PA is clearly associated with higher function and independence in older age (110).

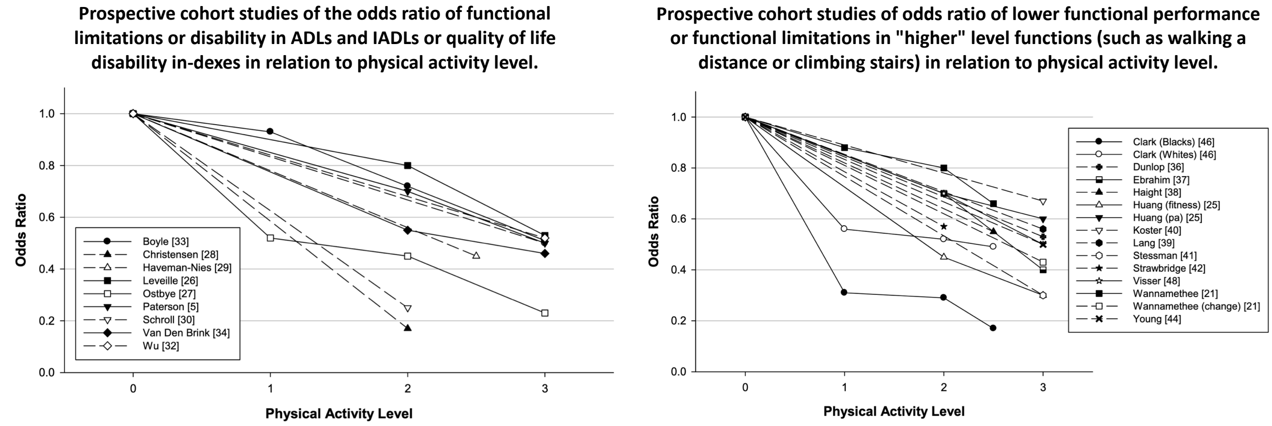


Figure 2. Prospective cohort studies included in Paterson et al. (110). The left figure illustrates the odds ratio of functional limitations or disability in activities of daily living or instrumental activities of daily living on physical activity levels from low to high, presented in each prospective cohort study. The right figure illustrates the odds ratio of lower functional performance or functional limitations in higher level functions such as walking a distance or climbing stairs on physical activity levels from low to high, presented in each prospective cohort study. The original figure in Paterson et al. are separated in two independent figures.

PA levels are not only associated with functional outcomes and independence. PA levels are also highly associated with mortality and pre-mature death in older adults (10,17,22,30). In sum, it is well established that PA levels are associated with being free from disease, avoiding morbidity, functional status, independence and thus the chance of achieving healthy ageing (2,3,17,22,110). In the recent years, research has even shown that older adults, including chronically ill patients, will gain longevity benefits and functional improvements by increasing the PA levels (3,21). Furthermore, even small changes have substantial clinical relevance (17,18,22) and differences of 1,000 daily steps have been reported to decrease the risk of all-cause mortality by 11% (22). In the end, Hippocrates was probably right when he introduced the notion that *exercise is medicine* around 400 years BCE, and it is now certain that it applies for older adults as well (111).

Global physical activity recommendations

The population of interest in the thesis is older adults in general. No specific subgroup (e.g. patients, high risk groups or frail individuals) is of interest and thus, the general recommendations for PA are relevant for the design and scope of the studies included.

The WHO recommends that older adults perform at least 150-300 minutes of MVPA or 75-150 minutes of VPA throughout the week for health benefits (29,112). Previously, the WHO recommendations also specified that MVPA should be undertaken in 10-minute bouts (29), but due to emerging evidence that also shorter bouts of PA have equal positive health benefits, the recent recommendations from 2020 do not include this specification (18,112,113). The recommended amounts of MVPA and VPA are considered the rule of thumb for most populations, and as such these recommendations also apply to older adults (though they also include further recommendations for muscle-strengthening activities, etc.) (29). MVPA is defined as PA with an intensity above three Metabolic Equivalent of Task (MET) and VPA is defined as a level of PA above six MET, where MET is defined as the ratio of the intensity a person is working relative to their resting metabolic rate (the energy consumption of sitting, equivalent to a consumption of 1 kcal per kg per hour) (114). Common MVPA activities include but are not limited to brisk walking, dancing, gardening and walking domestic animals for adults in general (114). For older adults it has been suggested that the intensity of regular walking has been misclassified and underestimated, and that older adults (especially impaired or slow walkers) do spend time in MVPA when walking at their usual pace (115,116). Thus, the global recommendations for PA specify not only the amount of PA but also the intensity, and older adults might achieve the intensity needed for health benefits solely by walking. This is supported indirectly by the findings of Lee et al. that it is the number of steps taken per day, and not the intensity or pace, which is significantly associated with mortality-rates (22).

As stated in the introduction, the prevalence of inactivity among European older adults is disturbingly high (10,12). At first glance, the rate of adherence to the global recommendations for PA is equally disturbing, as only one in four Danish older adults and one in eight British older adults are reported to meet the recommendations (14,117). The literature is inconclusive when it comes to how much older adults would need to walk to meet the global recommendations on PA, and it has been suggested that the recommendations are too high for the general population of older adults (118,119). New evidence reports substantial and clinically relevant health benefits in the form of large reductions (hazard ratio of 0.59) in mortality rates when comparing older adults with daily step counts of 2,700 with older adults who reach 4,400 daily steps (22). This is supported by other studies, and overall, there seems to be a plateau in risk reduction around 7,000 steps per day among older adults. Hence, it could be considered a relevant accommodation of the global PA recommendation for older populations (10,22,110,118,119).

In conclusion, walking is a free and common form of PA among older adults. Although the global recommendations on PA target MVPA and MET constructs, steps seem to work very well in public health research among older adults as a measure and construct of PA (17,22). Especially for self-monitoring among individuals, which is highly relevant to this thesis, steps have the advantage over MVPA or MET constructs that most consumer-available PAMs quantify steps, which will be elaborated in the following section.

Physical activity monitoring

When measuring PA, validity and precision are inversely associated with practicality and feasibility. Consequently, researchers will have to decide when, how and why they will measure PA to decide upon an instrument tool (120). For the sake of simplicity, this thesis distinguishes between four categories of PA measurement tools: energy expenditure estimation using doubly labelled water, mechanical pedometers, accelerometers and self-reported PA measurements.

Doubly labelled water, first used in 1955 by Lifson et al. (121), measures the disappearance rate of labelled isotopes from urine samples to estimate carbon dioxide production, which can be used to estimate the total energy expenditure (122,123). Doubly labelled water is highly accurate and considered the gold standard of energy expenditure measurements (79). However, it is time consuming, invasive and not applicable for large scale PA measurements or longer experimental designs (69,79,124). Furthermore, doubly labelled water measures energy expenditure, not PA. The two constructs may be highly correlated, but they are not the same (69). Energy expenditure might also be the most important construct when investigating inactivity, chronic disease status and health among other populations, but PA, as previously defined, is a more important construct among older adults because of its association with and key role in healthy aging.

Mechanical pedometers quantify PA through steps. Leonardo da Vinci is considered the original inventor of the mechanical pedometer, which was developed to count the steps taken by troops from the Roman military and thus help increase the precision of his maps (125). Centuries later, Thomas Jefferson introduced a more modern (smaller and more easily wearable) mechanical pedometer to the US, which had its roots in French and Swiss horology and worked through counting movements of a pendulum on a pocket watch like dial (126). By the late 19th century, the mechanical pedometers begin to appear in the health research literature, in studies on the amount of steps taken by medical professionals in hospitals (127). One of the first links between objectively measured PA and health were published by Larsen et al. in 1949, on the difference in PA levels between obese and non-obese patients, as cited in Chirico and Stunkard (128). In 1965 Dr. Yoshiro Hatano, concerned about the rise of obesity in Japan, used mechanical pedometers to investigate how many steps people should take on average to reverse this rise (129). Later, Dr. Hatano started selling a very simple and useful mechanical pedometer called the “Manpo-Kei” (10,000-steps-meter), illustrated in Figure 3 below, and became the first medical professional to incorporate goal-setting and results from mechanical pedometers in his practice (129). His 10,000 steps per day-goal is still commonly used in lay and scientific areas to this day, although it is recently coming under challenge from more recent evidence that suggests it may be an unachievable goal for many and that health benefits can occur at less than 10,000 steps per day (17,22,118). In sum, mechanical pedometers are the root of PA monitoring and have played an important role in both activity and epidemiological research, as well as in public health and behavioural change programs. However, as technology has progressed, modern step counters have instead adopted electronic accelerometery and software to measure the accumulation of steps, which offer greater reliability and validity than mechanical pedometers.

Figure 3. Manpo-kei (”10,000 steps meter”) marketed in Japan by Y. Hatano in 1965 (129).

Accelerometers use small crystal and a micro-electromechanical systems to measure accelerations in relation to gravity by converting the energy from the crystal into electronic signals (130). This allows accelerometers to produce raw acceleration counts in all three axes (130–132). Simply put, accelerations as a result of movement are directly quantified, while acceleration as a result of gravitational force provides the data on the orientation of the accelerometer, which is crucial for identifying different types of movement (130–132). Each type of accelerometer provides different raw accelerometer counts and as human movement is highly complex, the hardware and software need to be calibrated to each measurement purpose and the location of the accelerometer. From the raw accelerometer counts, desirable parameters and constructs can be calculated and quantified, such as time spent in MVPA, PA with different intensities (e.g. MET-minutes) and sedentary activities, but also specific body movements like moving from sitting to standing positions, jumping, squatting and – crucially to this thesis – walking (130–132). Research-grade accelerometers like Actigraphs (https://actigraphcorp.com/actigraph-wgt3x-bt/) or Sens Motion (https://sens.dk/) are considered the gold standard of PA measurement (69,79,130), but consumer-available accelerometers have seen greater use in recent years (67). Consumer-available accelerometers, which include the likes of Fitbits, Garmin Fitness Trackers, Jawbones and other smart watches, do not allow individual users to modify the software and algorithms used for activity pattern recognition, but are more readily available and less expensive than research-grade accelerometers, but most importantly, they provide real-time feedback on PA to the user (35,67). In general, consumer-available PAMs are considered to provide valid and reliable measurements of steps, while they tend to overestimate other parameters, such as MVPA, compared to research-grade accelerometers (78,133,134). In summary, research-grade accelerometers may be considered the gold standard of PAMs, but consumer-available accelerometers have the potential to serve both the facilitator and driver of behavioural change while simultaneously quantifying PA levels.

In the original studies conducted in the thesis (**study II, study III** and **study IV**), objectively measured PA and self-reported PA were used as outcomes and constructs for validation purposes. The studies used consumer-available PAMs deemed appropriate for the population of older adults (because of their long-lasting battery life and simple design) as objective PA measurement tools. The Nokia GO, Jawbone Move UP, Garmin Vivofit 3 and the Misfit Shine were all included for criterion validation, and the Garmin Vivofit 3 was additionally used as an outcome measure in **study III** and **study IV**.

The Garmin Vivofit 3 tri-axial accelerometer is commercially available and had not been validated among older adults before **study II** (135). The Garmin Vivofit 3 can be fastened in a belt-clip or with a wrist band. It operates on a button cell battery and does not need to have its battery changed for six to 12 months of daily use. It has enough built-in memory for four weeks of activity data and connects to a smartphone using Bluetooth. Garmin Vivofit 3 provides feedback on steps (measured by the accelerometer), energy expenditure (burned calories calculated from step volume and intensity) and feedback on daily goals according to either of these. After synchronising with the Garmin Connect Application (https://connect.garmin.com/), daily goals can be modified and adjusted manually or automatically.

Figure 4. The Garmin Vivofit 3 tri-axial accelerometer.

Despite their good accuracy, validity and general measurement properties, PAMs have the limitation that they only measure the PA when worn. Studies are thus encouraged to include wear time in the analysis (82). A commonly used alternative to objective measurement of PA are participant-reported measures, such as PAQs or diaries, which cost-effective and feasible methods (83). Because of their practicality, PAQs are especially widespread in public health research and in experimental study designs. However, their measurement properties suffer from social desirability and recall bias (84–87). A systematic review of PAQs used in populations with older adults identified 40 different instruments and concluded that, due to large inconsistencies in how older adults understand and report PA, future research should consider the participants' conceptual understanding of PA and seek to develop questionnaires that inquire more clearly about functional, light-intensity PA as well as sedentary activities (136). No existing PAQ seems ideal for all types of older adults (136,137) and researchers must balance the need for population-specific questionnaires with the need to generalise to other populations, when more general questionnaires are needed. Furthermore, even though there are several PAQs that are specifically designed for use among older adults (137), all of them display large measurement error and only the Physical Activity Scale for Elderly instrument and the Physical Activity and Sedentary Behaviour Questionnaire show adequate reliability and construct validity (136). Unfortunately, these two PAQs that were specifically designed for older adults have not been translated to Danish and validated in a Danish context. Furthermore, it is recommended that researchers use, interpret and improve future existing PAQs instead of developing new ones (136).

Study III and **study IV** used the IPAQ-SF and the NPAQ-short as self-reported PA measures. The International Physical Activity Questionnaire was originally developed in 1998 by an international consensus group in Geneva to allow for large-scale global assessments of PA (89). Since then, the International Physical Activity Questionnaire has become the most widely used PAQ. It is available in two versions: A 31-item long form and the nine-item short form, which was used in this thesis (89,138). The IPAQ-SF quantifies the amount of 'MVPA', 'vigorous PA' (VPA), 'MET-minutes', 'walking time' and 'sedentary time' performed in the last seven days. It allows three PA categories to be derived from the responses: 'low activity level', 'moderate activity level' and 'high activity level' (89). The IPAQ-SF has been reported to have acceptable test-retest reliability (89) and low-to-moderate concurrent validity compared with objectively measured PA (86). The Danish version of the IPAQ-SF has previously been used among older adults (139). The Nordic Physical Activity Questionnaire is a Danish survey tool based on telephone interviews originally designed to assess MVPA and compliance with PA recommendations (140). The NPAQ-Short is a revised version with two items (90). Besides quantifying MVPA, the questionnaire yields four categories of PA: 'inactive', 'insufficiently physically active', 'sufficiently physically active' and 'optimally physically active'. 'Inactive', 'insufficient physically active' are further combined to report a binary category of adherence or non-adherence to WHO recommendations on PA. The NPAQ-Short has only been used in research for a few years, but it has been shown to have moderate validity (correlation with objectively measured PA) in a Danish population of adults (90).

Validity of physical monitoring

In **study II** and **study III**, this thesis investigated the validity of PA monitoring. **Study II** investigated criterion validity between PAMs and visually counted steps, while **study III** investigated concurrent validity of the PAQ's ability to reflect objectively measured daily steps. Criterion validity is a type of measurement validity that determines how a score correlates with the score of the criterion test, while concurrent validity is a sub-category of criterion validity that determines how well the score of a test reflects or correlates with the score of another test measured at the same time, within the same construct (141,142). In summary, criterion validity was used in **study II** because visually counted steps were seen as the key criterion for the PAMs. Concurrent validity was used in **study III** as a more limited type of validity because the sub-categories of the PAQs did not produce daily step counts, but rather MVPA, MET-minutes and walking time, which were expected to correlate with and reflect the daily step counts.

Behavioural change

As mentioned in the introduction, one in four older adults in Denmark above the age of 65 and one in two above the age of 75 do not meet the WHO recommendations for minimum PA (13,14). Behavioural change related to PA has been of interest since Hippocrates, the father of scientific medicine, who was the first physician to prescribe exercise to patients (111,143). Countless behavioural change theories and ways of motivating individuals to act a certain way have been proposed and used through time. The thesis investigates how feedback from PAMs can be used to facilitate healthy behavioural change in older adults, and as the project evolved, how MI can be used to increase the effect of PAMs. Thus, this section includes the rationale and evidence for using PAMs to enhance PA levels in general and specifically in older adults, as well as introducing motivational interviewing as a theoretical and practical approach to PA interventions.

Physical activity monitoring to facilitate behavioural change

The growing trend toward self-monitoring among individuals, also known as "the quantified self," and the rise of consumer-available PAMs make PA monitoring, as well as sleep and diet monitoring, an interesting and growing new area of research (32). It is also undoubtedly a growing field, as Figure 5 below illustrates.

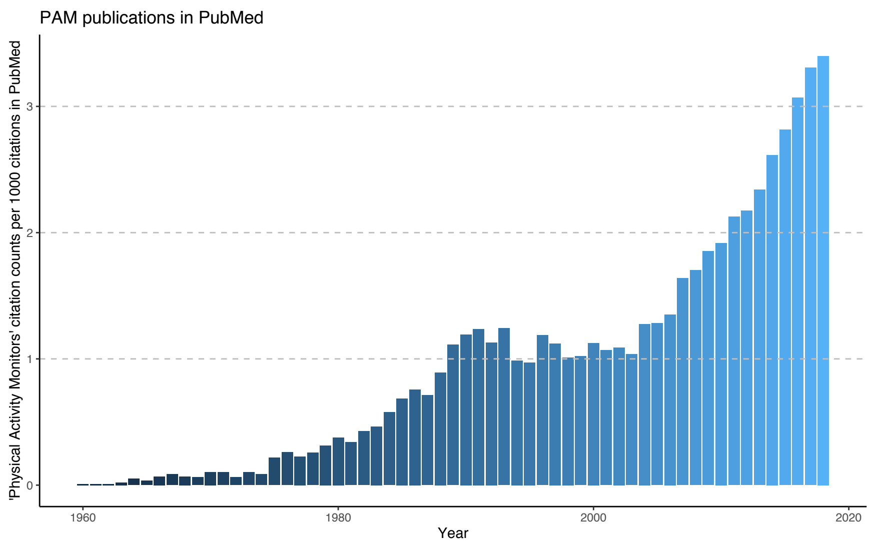


Figure 5. "Physical Activity Monitor" citation counts per 1000 citations in PubMed per year since 1960.

The effectiveness of PAM-based interventions in terms of PA and sedentary time have already been investigated in adults (36,144). However, their effectiveness might be different in older adults and the literature seems inconclusive, as some studies have reported PAM-based interventions to be more effective than both passive and active control interventions (37–42), while others have not (43–47). No systematic review of the literature had been conducted or published before **study I** of the thesis, which fills the evidence gap of PAM-based interventions among older adults.

Motivational interviewing to facilitate behavioural change

Study I concluded that no further passive comparisons should be pursued in relation to PAM-based interventions, and that future studies should either investigate superiority trials comparing PAM-based interventions to other behavioural change interventions or investigate whether the effect of PAMs could be increased by add-on interventions. The latter approach was chosen for **study IV** of the thesis.

MI is "*a directive client-centred counselling style for eliciting behaviour change by helping clients to explore and resolve ambivalence*." According to the founders of the approach, Rollnick and *Miller, "It is most centrally defined not by technique but by its spirit as a facilitative style for interpersonal relationship"* (145). MI evolved as a counselling style for problem drinkers, first described in 1983 (146), and is based on the following principles: 1) ambivalence about behaviour change is normal and an important motivational obstacle, 2) ambivalence can be resolved by working actively with the client's intrinsic motivations and values, 3) the relationship between the counsellor and the client is an alliance and a collaborative partnership to which both partners bring important knowledge and expertise, and 4) empathic and supportive, yet directive, counselling is encouraged over direct argumentation and aggressive confrontation that may increase client defensiveness and reduce the likelihood of behavioural change (147). Even though MI was originally developed for problem drinkers, it has been used in relation to PA and inactivity in several research projects (148,149). The Transtheoretical Model is the underlying theory of behavioural change in MI (147,150). Inspired by the work of Eakin et al. and Ismail et al., **study IV** also drew on Social Cognitive Theory to support the counsellors by addressing and operationalising the constructs (151–153).

The Transtheoretical Model explains behaviour change as following a series of stages: 1) precontemplation, 2) contemplation, 3) preparation, 4) action, and 5) maintenance (154–156). In MI, this theoretical approach is used to guide the course of counselling. In the precontemplation stage, the participant could receive guidance on the benefits of the behavioural change and risks of no change; in this case, the association between PA and healthy ageing. In the contemplation phase, barriers and misconceptions should be addressed and support systems identified. In the preparation phase, realistic goals and timelines for changes should be discussed to provide positive reinforcement, in line with the strategy in the action phase, where the counsellor's primary focus is to provide positive reinforcement. In the maintenance phase, encouragement and support should be provided to prevent relapse to a lower level of PA (154–156). Social Cognitive Theory explains human behaviour from cognitive/personal factors, environmental factors and behavioural factors (153). Consequently, the MI-counselling used in **study IV** sought to facilitate behavioural change by targeting the cognitive factors. Expected PA outcomes were managed and guidance was provided about the relationship between PA and health in older age. Environmental factors were targeted by the counsellor advocating and discussing social networks for PA and peer support, and behavioural factors were targeted through increased self-efficacy in the participants. To have a valid and responsive measure of change, we investigated, translated, and validated the Self-Efficacy for Exercise Scale (SEE) and the Outcome Expectancy for Exercise Scale (OEE) in older adults as a parallel project to the thesis (paper forthcoming) (157–159). In sum, the MI counselling used in **study IV** actively applied the Transtheoretical Model and the Stages of Change to navigate the conversations with the participants, and Social Cognitive Theory to further guide the counselling and provide useful constructs for counsellors.

MI was chosen as an add-on intervention in **study IV,** as it is commonly used and hence well-established in health and social systems in Denmark (57–63). Specifically with older adults and PA, MI has been reported to increase PA levels among patients with hip fracture (65) and patients heart failure (64). Furthermore, the combination of MI and PAMs has been reported to acceptable by older adults (66).

Evidence-based research

This section provides the philosophical and academic foundation of how best-practice evidence can be produced to answer the aim and research questions of the thesis. As this section will clarify, a great deal of research is wasted on what is already known, and academics should be responsible and consistent in their production of new research (160–162).

The present PhD thesis sought to investigate how PA monitoring might be used to enhance PA levels among older adults. This aim could have been answered in numerous ways and consequently elicited several different study-aims and publications. According to Chalmers and Glasziou, 85% of new research is wasted, at an estimated annual cost of 149 billion euro globally (162). To avoid contributing to this waste of research and funds, the aim and order of the studies conducted in the course of this PhD project sought to follow the philosophy of evidence-based research (160). To avoid bias and to secure publication of the results, the studies were conducted so to adhere as rigorously as possible to current best practice guidelines (163–166). Evidence-based research simply means that no new research should be done without a systematic review of the current literature, and that research that fails to do so is unethical, unscientific and wasteful (160). Juhl and Lund (167), through the consensus presented in BMJ by Garner et al (168), argues that before conducting a systematic review, the research question should still be clinically important, unanswered, and that the methods of the planned systematic review should follow a structured and pre-planned protocol that defines the research question, inclusion criteria, search methods, selection procedures, quality assessment, data extraction, and data analysis (167).

A proportion of the mentioned waste of research is caused by lack of methodological quality in clinical trials (169). Though the quality of clinical trials have increased, especially after the introduction of the Consolidated Standards of Reporting Trials (CONSORT) statement (163), absence of trial protocols, low statistical precision (power), arbitrary analyses, and extrapolation on random extreme findings are still common (170). Bandholm et al. have published the PREPARE trial guide for RCTs that addresses several aspects of the planning and preparation of clinical trials (161). One aspect includes a well-framed and good research question relevant to all stakeholders (patients, clinicians, decision makers and others), which must follow a novel and relevant new systematic review of the current literature to follow the philosophy of evidence-based research (160,161,171), as outlined above. Other aspects of the PREPARE checklist (161) include, but are not limited to, choosing the right outcomes (according to the Core Outcome Measures in Effectiveness Trials (COMET) and Consensus-based Standards for selection of health Measurement Instruments (COSMIN) initiatives) (142,172), pre-defining the statistical analysis plan and the handling of missing data, and finally registering and publishing the trial protocol.

Considerations for the methods and projects in the thesis

Balancing the need to adhere to best-practice and gold standard recommendations, as mentioned above (160,161), with a tight and time-constrained project, limited funding and opposing interests from stakeholders can be difficult. To follow the recommendations from evidence based research (160), I chose to begin my PhD project with a systematic review (**study I)** that investigated whether PAMs had already been shown to be effective in older adults and to investigate the heterogeneity of the results to clarify if some sub-populations of older adults had different effect sizes than others.

To be as transparent as possible and to follow the best practice recommendations, the protocol for **study I**, the systematic review, was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) statement (173). It was also registered in The International Prospective Register of Systematic Reviews (PROSPERO) database before the search was conducted, and it was subsequently published in BMC Systematic Reviews journal (paper A). The results of the systematic review (paper B) were according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA) statement (164) and followed the recommendations of the Cochrane handbook (165) as closely and rigorously as possible without being conducted as an actual Cochrane review, due to time constraints. To avoid research waste (162), the hypothesis, design, and preparation of the experimental **study IV** evolved directly from the results of **study I**. Furthermore, workshops with older adults were conducted at activity centres in Copenhagen to allow the perspectives of potential participants to be directly incorporated in the design of **study IV.** MI was selected as a relevant, already well-defined and implemented behavioural change intervention. A core set of outcomes was chosen to investigate the expected behavioural change more broadly. To ensure the relevance of the included outcomes, three validation projects were conducted before **study IV** commenced: 1) translation and validation of the SEE scale and OEE scale (not included in this PhD thesis), 2) validation of consumer-available PAMs among older adults (**study II**), and 3) validation of the two PAQs chosen to support the primary outcome measurement (**study III**)**.** Finally, I chose to publish the protocol (paper E) for the RCT (**study IV**) as a priority. In sum, several factors affected my choice of studies and methods for the thesis throughout the three years, and the adaptability of my research approach has been maintained by having my early studies guide the final studies. It was ultimately not possible to follow every recommendation of the best-practice guidelines (142,160,163–165), and thus this project comes with several limitations, which will be discussed later. Nevertheless, I sought to conduct evidence-based research from the beginning through to the end of the project (160).

Summary of existing evidence

In 2007, a highly cited systematic review in the JAMA by Bravata et al. suggested that PAMs could be used effectively to motivate and facilitate increased PA levels among adults (36). However, as older adults might react differently to such interventions, the results from the review could not be extrapolated to older adults. Several RCTs have reported PAM-based interventions to be effective in older adults (37–42), but others have reported them to be ineffective (43–47). The above-mentioned studies showed a great deal of large clinical, methodological and statistical heterogeneity, and consequently demonstrated the need for a systematic review of the literature on the effectiveness of PAM-based, PA-promoting interventions aimed at populations of older adults.

Consumer-available PAMs are more now frequently used than ever among the general population (35,67), but the measurement properties of specific PAMs should be evaluated before they can be used as clinical outcome measurement instruments. The algorithms underpinning most consumer-available PAMs are developed from younger populations (33), but as stride length, gait speed and assistive devices have been found to affect the validity of PAMs (70,71), the validity of consumer-available PAMs would need to be investigated among older adults specifically. Few studies have examined their validity among older adults (71–78), and no investigation of PAM placement has been conducted using the same devices. Hence, it would appear necessary to conduct validation of each PAMs of interest before they are used as outcome measurements among older adults.

PAQs are easy to use and feasible in most situations. However, their psychometric properties are limited by recall and social desirability biases (84–87). The IPAQ-SF (88,89) and the NPAQ-short (14,90) are two relevant PAQs for intervention studies. Unfortunately, the criterion validity of these instruments remains unclear among older adults and would need to be investigated before the two PAQs could be used or their results taken as conclusive for older adults (86,91,92).

MI is commonly used in health and social systems in Denmark (57–63), and has been reported to increase PA levels among patient sub-groups of older adults (64,65) and to be acceptable in combination with PAMs (66). As such, MI represents a highly relevant behavioural change theory and intervention to investigate as a potential add-on to increase the expected effect of PAMs within a population of Danish older adults.

Contributions of this thesis

As mentioned in the section above, there is an extant evidence gap regarding 1) whether PAMs are effective in increasing the PA levels among older adults, 2) whether specific consumer-available PAMs are a valid measure of PA in older adults and if the placement of the specific devices affects the results, 3) whether the IPAQ-SF and NPAQ-Short are sufficiently valid to be used in populations of Danish older adults, and finally 4) whether the expected effect of PAMs can be increased by adding MI to a PAM-based intervention. This current PhD project investigated each of the above-mentioned evidence gaps, and thus addressed the overall question of how PA monitoring can be used to enhance PA levels among older adults.

1) **Study I** found a low quality of evidence for a moderate effect of PAMs in increasing the level of PA among older adults. Furthermore, it was clear that no future studies comparing PAMs with passive control interventions were needed.
2) **Study II** validated three different consumer-available PAMs that were specifically relevant for older adults. The Garmin Vivofit 3 was the best performer. It further found that wrist-worn PAMs should not be trusted in a population of older adults walking freely with older adults using rollators.
3) **Study III** found both that neither IPAQ-SF nor the NPAQ-short reflect the criterion measure of PA adequately. The results cannot support the use of the PAQs among older adults.
4) **Study IV** found a clinically relevant, but not significant, effect of adding MI to a PAM-based intervention. Unfortunately, due to a small sample size and lack of power, future studies would need to reproduce this finding before decision-makers can combine MI and PAMs in preventive or rehabilitation programs among older adults.

In summary, the thesis investigated how PAMs could be used to increase the daily PA level among older adults. Consumer-available PAMs do have the potential to serve as valid outcome measurements and effective intervention facilitators to increase PA levels. PAQs should be used with the greatest caution among older adults, however, and should not be trusted alone. MI might increase the effect of PAMs among older adults, but future studies are warranted.

MATERIAL AND METHODS

The four aims of the thesis have been investigated through four individual studies and six peer-reviewed papers, five of which have been published and one which has been submitted, but was under review at the time of this project's submission (135,174–177). Figure 6 below illustrates that **study I** was initiated and conducted to inform the other studies, especially the aim and design of **study IV**. **Study II** and **study III** were conducted to validate the outcome measurement instruments used in **study IV**.

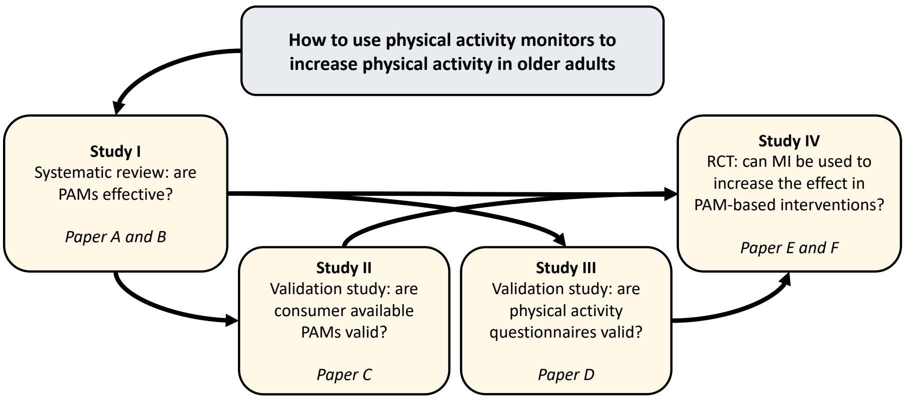


Figure 6. Illustration of the rationale, flow and methods of the thesis. Chronologically, **study I** was initiated in the beginning of the process to investigate whether physical activity monitors already had been shown to be effective among older adults and to inform the other studies. **Study II** was conducted after the main results of **study I** had been interpreted to inform the design of **study IV**. **Study III** and **study IV** use data from the same participants. The results of **study III** are relevant to interpretation of the findings in **study IV**.

This section presents the material and methods of the thesis and the four individual studies. As **study I** was designed as a systematic review of existing literature, it constitutes secondary literature by nature and will be described in this section independently. **Study II, study III and study IV** were original research investigations and will be described together in this section. Below, a '*methods at a glance table'* provides a brief overview of the methods used in the four studies.

Table 1. Methods at a glance.

Study # and objective	Design	Methods
Study I **To estimate the effect of PAM-based interventions on PA behaviour in participants aged 65 and above.**	Systematic review and meta-analysis investigating the effect of feedback from physical activity monitoring among older adults.	Searches in five databases were performed. RCTs and randomised cross-over trials were included if they compared interventions where the participants received feedback from PAMs with control interventions where the participants did not receive any feedback. Random-effects meta-analysis using Hedges' g were used to pool the study results. Results on physical activity, sedentary time, MVPA, physical capacity, body mass index and HRQoL.
Study II **To investigate the criterion validity of four consumer-grade PAMs in older adults and to investigate whether the measurement properties were affected by placement and assistive devices.**	Validity study investigating the criterion validity of four PAMs found relevant to older adults.	Community-dwelling older adults above the age of 65 with the ability to walk independently with or without assistive devices performed self-paced walking while their steps were visually counted. The participants wore 16 monitors (four from each device; Misfit Shine, Nokia GO, Jawbone UP Move and Garmin Vivofit 3). ICC (1,2) were used to express criterion validity.
Study III **To investigate the concurrent validity between the IPAQ-SF and the NPAQ-Short and objectively measured daily steps among older adults.**	Validity study investigating concurrent validity between two PAQs and objectively measured daily steps.	Community-dwelling older adults above 70 years of age included in study IV were eligible for inclusion if they had more than four valid wear days in the baseline week and answered the baseline survey with the PAQs. Spearman's rho was used to investigate concurrent validity.
Study IV **To investigate the effect of MI as an add-on intervention to a PAM-based intervention measured in community-dwelling older adults.**	Two-arm parallel group RCT investigating if a PAM+MI intervention were more effective than a PAM intervention	Community-dwelling older adults above 70 years of age were randomised to a control intervention (PAM) or an experimental intervention (PAM+MI) where MI were added to a PAM intervention. Average daily step count, was analysed following the ITT principle with multiple imputations.

Abbreviations: Physical Activity Monitor (PAM), Randomised Controlled Trial (RCT), Standardised Mean Difference (SMD), Health-Related Quality of Life (HRQoL), Interclass Correlation Coefficient (ICC), International Physical Activity Questionnaire Short Form (IPAQ-SF), Nordic Physical Activity Questionnaire Short (NPAQ-Short), Physical Activity Questionnaire (PAQ), Motivational Interviewing and Physical Activity Monitoring (MIPAM), Motivational Interviewing (MI), Intention-To-Treat (ITT), 95% Confidence Interval (95%CI), University of California, Los Angeles (UCLA).

Material and methods of study I

Eligibility, searches and data extraction

The systematic review and meta-analysis was conducted according to the recommendations of the Cochrane Handbook (165) and reported according to the PRISMA statement (164,173). The protocol was pre-registered in the PROSPERO database (CRD42018083648) before the systematic search was performed.

RCTs were considered eligible for inclusion in this systematic review if they compared PAM-based interventions where the participants received feedback from PAMs with control interventions where participants did not receive feedback. The mean age of the participants in the studies had to be above 65 years, and the participants in the studies had to walk independently.

The primary outcome, daily PA, was extracted from the included studies, favouring objectively measured PA. The secondary outcomes included sedentary time, MVPA, physical capacity, Body Mass Index (BMI), and HRQoL. Outcomes were extracted at baseline and end-of-treatment if possible.

The final search was performed on April 26, 2018 in the following electronic bibliographic databases: the Cochrane Controlled Register of Trials, CINAHL, SPORTDiscus, EMBASE and MEDLINE (PubMed). The clinical trial register www.clinicaltrials.gov was searched to identify ongoing studies. The search strings consisted of a combination of keywords and subject headings for 1) older adults, 2) PAMs, and 3) RCT designs using the Cochrane's highly sensitive filter for identifying RCTs (178). The search strings are included in the study protocol (174), alongside the search matrix.

Two reviewers independently screened title and abstracts as well as full text reports to exclude irrelevant study reports. Inconsistencies were solved through discussion or inclusion of a third reviewer. The same two reviewers extracted data using a pre-determined data extraction form independently and assessed risk of bias in individual studies for each outcome using the RoB 2.0 tool (179).

Summary measures and additional analyses

Intervention effects were expressed as Standardised Mean Differences (SMD) at 95% CI because different outcomes were used in the included studies. For interpretation purposes, the SMDs were translated to absolute effects by multiplying the value with SDs from the largest study available with an overall low risk of bias. Adverse events and withdrawals were expressed as relative risks with 95% CIs.

As the assumption of one true effect size in all studies could not be met, a random-effects meta-analysis adjusted to Hedges' g was performed (180,181). As recommended, the analysis was performed using post-intervention scores only (182). The I^2 statistic was used to determine any between study variance not explained by chance (183). Random-effects meta-analyses sorted by subgroups were used to investigate heterogeneity with categorical variables and univariate meta-regression analysis was used to investigate heterogeneity with continuous variables. Publication bias was investigated using the Eggers test, where positive results are adjusted using the Duval and Tweedie nonparametric "trim and fill" adjustment (184). StataCorp. 2017 *(Stata Statistical Software: Release 15.* College Station, TX: StataCorp LLC) was used for all statistical analyses. An alpha level of 0.05 was set as the threshold for statistical significance. For further detailed description of materials and methods of **study I**, please see the study protocol (174) and the final study paper (175).

Material and methods of study II, Study III, and study IV

Ethics

Study II was approved by the Danish Ethics Committee (journal number: H-17033310) and oral and written informed consent forms were completed before participants were included in the study. **Study III and study IV** were conducted with waivers from the Danish Ethics Committee of the Capital Region of Denmark (Journal no.: 18004960). The Danish Data Protection Agency Approved the handling of data (Reference no.: 514-0268/18-3000). Informed consent from the participants was collected electronically via *SurveyXact* (185). Prior to agreeing and signing the consent survey, the participant received written information about the study.

Participants

The inclusion criteria for participants differ slightly between **study II, study III** and **study IV**. For **study II,** older adults from five activity centres in Copenhagen were recruited. Participants were considered eligible if they were above 65 years of age, community-dwelling and able to walk independently with or without walking aids. **Study III** used baseline data (before randomisation) for participants enrolled in **study IV**. Participants were considered eligible for inclusion in **study IV** if they were community-dwelling older adults above the age of 70, owned a smartphone able to install the *Garmin Connect* application*,* had an e-mail address, were able to complete the study survey, and were able to receive a telephone interview. Exclusion criteria included: 1) cognitive impairment from moderate to severe dementia or Alzheimer's disease, 2) active chemotherapy or palliative care for cancer, and 3) major mobility impairment. Furthermore, to be considered eligible for inclusion in **study III,** the participants had to provide daily step data for at least four days and should have completed the electronic survey.

Design and procedures for ***study II and study III***

To investigate the criterion validity of relevant PAMs, participants in **study II** performed self-paced walking for six minutes on an unobstructed flat track. Four consumer-grade PAMs were considered eligible for inclusion as they could be fastened at the hip as well as on the wrist, were simple in function and design (requiring no technical skills to be operated), provided real time feedback and were powered by a button cell battery with a long battery life. PAMs were fitted at each wrist and each hip of the participants (16 monitors in total). Two physiotherapists, blinded from each other's counting, counted the number of steps using a click-counter.

To investigate the concurrent validity of the IPAQ-SF and the NPAQ-Short questionnaires, participants from **study III** received a Garmin Vivofit 3 PAM by mail and were asked to wear it during all waking hours for one full week. After wearing the PAM for one week, the participants received a link to the online study survey containing the two PAQs. The IPAQ-SF consist of nine items and the following quantified constructs were used as outcomes in **study III**: 'VPA', 'MPA', 'MVPA', 'MET-minutes', 'walking time' and 'sedentary time'. Three PA categories were derived from the responses: 'low activity level', 'moderate activity level' and 'high activity level' (89). The two-item NPAQ-short (90) was developed to measure MVPA and adherence to WHO recommendations (90,140). The following quantified constructs were used as outcomes in **study III**: 'VPA', 'MPA', and 'MVPA'. From this, four PA categories can be calculated: 'inactive', 'not sufficiently physically active', 'sufficiently physically active', and 'optimally physically active'. Furthermore, NPAQ-short categorisation of adherence to the WHO recommendations was used. Further description of the two PAQs can be found in the background section.

*Design and procedures for **study IV***

Study IV was conducted as a 12-week, investigator-blinded, two-arm, parallel-group superiority randomised controlled trial. The only change between the published study report and the protocol (177) was a failure to reach the sample of 128 participants that had been desired in order to significantly show a moderate effect difference (half a standard deviation (SD)) with 80% power.

The control (PAM) group received a PAM-based PA-promoting intervention and the intervention (PAM+MI) group received the PAM-based PA-promoting intervention as well as an MI-intervention as an add-on intervention. Participants from both groups received a hip-worn Garmin Vivofit 3 to use during all walking hours in the intervention period and a pamphlet with national recommendations on PA for older adults. The PAM was linked to a *Garmin Connect* account with automatically adjusted daily goal setting. During the 12-week intervention period, the participants in the PAM+MI group were scheduled to receive seven telephone calls from trained and certified MI-counsellors. In brief, MI is a client-centred counselling style for eliciting behaviour change (145). The Transtheoretical Model is the underlying theory of behavioural change in MI and is used to explain behaviour change as following a series of stages: 1) precontemplation, 2) contemplation, 3) preparation, 4) action, and 5) maintenance (154–156). MI is commonly used and well-established in Denmark (57–63) and the detailed description and theory behind the intervention has been described previously in the background section of the thesis, the trial protocol and the full study report (177).

Fidelity of **study IV** was investigated by recording MI-sessions on a regular basis to provide counsellor feedback, with the participants' verbal consent. Based on a review of these recordings, a random segment of 20 minutes was selected for rating with the Motivational Interviewing Treatment Integrity Scale version 4 (186). The scale includes four global ratings, Cultivating Change talk, Softening Sustain Talk, Partnership, and Empathy (186). A median global score in each global rating domain of 4 and a Reflection to Question ratio of>1 was considered decent MI proficiency.

The primary outcome measure for **study IV** was the average number of steps in the intervention period. The hip-worn Garmin Vivofit 3 PAM was used to continuously and objectively measure the number of steps taken. The secondary outcomes included self-reported measures of PA, HRQoL, loneliness, and self-efficacy and outcome-expectancy for exercise. The instruments IPAQ-SF (187–190), NPAQ-short (91,191), The 5-level EuroQol-5 Domain (EQ-5D-5L) Quality of Life (QoL) questionnaire (192–195), UCLA Loneliness Scale (196,197), SEE (159), and OEE (158,198,199), and the Copenhagen Social Relations Questionnaire (200,201) were used as secondary outcome measures and are described in greater detail in the study protocol (177). Secondary outcomes were collected at baseline and at the end of treatment.

Included participants were randomly assigned to the two groups with a 1:1 allocation ratio. The participants were randomised in blocks of minimum four, stratified by average step count and sex. Due to the nature of the behaviour change intervention, participants and interviewers were not blinded to allocation but the principal investigator was. Information about data collection and management can be found in the study protocol (177).

*Statistical methods for **study II, study III** and **study IV.***

For all the three studies Quantile-Quantile plots were used to evaluate the distributions of standardised residuals for continuous data. Variables determined normally distributed were summarized with means and 95% Confidence Intervals (95%CI) or SDs. Variables considered non-normally distributed were summarised with medians and interquartile ranges. Categorical data were summarised as frequencies and percentage of the samples.

In **study II,** the criterion was the actual number of steps taken, defined as the average of the visually counted steps from the two testers. Data points were excluded when the PAM did not count any steps. The two-way random-effects interclass correlation model (ICC2,1) was chosen to assess the criterion validity (202). Hence, the absolute agreement between the PAM and the criterion is reported. ICC (2,1) values of less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and equal to or more than 0.90, were interpreted to indicate, respectively, poor, moderate, good, and excellent criterion validity. *A priori*, moderate criterion validity for all participants, poor criterion validity for participants with rollators and good criterion validity for participants without rollators was expected for all PAMs. Scatter plots with a generalised linear logit link model were used to illustrate absolute measurement error for each PAM.

Study III used Spearman's rank correlation coefficients between PAQ scores and daily steps to measure concurrent validity. Spearman's rank correlation coefficients (ρ) of less than 0.25, between 0.25 and 0.5, between 0.5 and 0.75, and equal to or more than 0.75 were interpreted, respectively, as the PAQs having little or no correlation, fair correlation, moderate-to-good correlation, and good-to-excellent correlation with daily steps. *A priori*, IPAQ-SF and NPAQ-short scores were hypothesised to have at least a fair correlation with daily steps. The IPAQ-SF 'Walking' and NPAQ-short 'Moderate Physical Activity' questionnaires were expected to have a moderate-to-good correlation with daily steps, as the description of the specific questions include walking. 'IPAQ-SF Moderate Activity' was expected to have little or no correlation with daily steps as the description of the specific question excludes walking (141). Concurrent validity of the PAQ categories (relationship with daily steps) were investigated with one-way analysis of variance models and unpaired t-tests.

Study IV analysed primary and secondary outcomes with multiple regression models, adjusted for sex, baseline daily step count and the baseline value of the specific outcome. In calculating the average daily step count, days with less than 100 steps were excluded as "days of non-wear." To follow the intention-to treat-principle, a Gaussian normal regression method (predictive mean matching with five runs and 50 iterations) was used to conduct multiple imputation using the baseline step count, gender and age in samples with averages from less than seven days of wear in the intervention period. To avoid inspecting the data before choosing an imputation, the first imputation run was used. Harms were evaluated by calculating the between-group relative risk for serious and non-serious adverse events. For illustration purposes, specifically for use in the thesis, p-values from the between-group difference from multiple regression analyses on daily PA were extracted from 5,000 multiple imputation runs (5 iterations each) and the distribution of p-values was plotted in a density plot. The RStudio MICE package was used to perform the multiple imputations and the PURRR package were used to perform the functional programming (loops).

To allow for a deeper discussion of study results, several post-hoc sensitivity analyses and illustrations were conducted for the thesis. The sensitivity analyses include unpaired Student's t-test to compare the age of participants with and without imputed data in **study IV**, a linear regression model for analysing the relationship between number of missing days and age in **study IV**, a Wilcoxon-test testing difference in number of missing days between participants with and without adverse events in **study IV**, and analyses without multiple imputations (complete-case and last observation carried forward) with similar models to the primary analysis in **study IV**. Post-hoc illustrations include a histogram with a density curve of the average step count from the participants included in **study III,** a distribution of extracted p-values for 5,000 multiple imputations for the primary analysis in **study IV**, as described above, scatter plots with best fitted lines and 95% CIs illustrating the association between secondary outcomes and average daily steps as well as a plot illustrating the association between age and number of missing days in **study IV.**

StataCorp. 2017, *Stata Statistical Software: Release 15.* College Station, TX: StataCorp LLC was used for all statistical analyses in **study II**. RStudio version 1.1.463 (2016, Integrated Development for R. RStudio, Inc., Boston, MA URL) was used for all statistical analyses in **study III** and **study IV.** An alpha level of 0.05 was set as the threshold for statistical significance in all studies.

RESULTS

To provide an accessible overview, the first paragraph presents a narrative summary of the most important results. The following four sections then provide more detailed results from each of the four studies.

Narrative summary of the results

Study I identified 21 RCTs with 2,783 participants, which together indicated the existence of low quality of evidence for a moderate effect of PAMs on daily PA, which translates to 1,297 more steps per day on average in the intervention groups. Subsequently, a moderate quality of evidence was found for a small effect of PAMs on MVPA time, which translates to 5.5 minutes more MVPA time per day on average in the intervention groups. There were no significant differences between intervention groups and control groups in the pooled results on time spent sedentary, physical capacity, BMI or self-reported HRQoL.

In terms of validity of the consumer-grade PAMs, four different devices were found to be relevant in an elderly population and were included in **study II**. In total, 103 older adults participated and were fitted with monitors to use while walking. The Nokia GO was excluded due to technical issues. The Jawbone UP Move had higher rate of excluded data due to technical issues and the wrist-worn monitors were also found to have a higher degree of excluded data, compared to the hip-worn PAMs. The hip-worn PAMs did not differ significantly in terms of measurement error or validity. The wrist-worn PAMs were found inadequate in measuring number of steps among older adults with and without rollators.

Study III investigated the validity of two electronically administered PAQs and found both the IPAQ-SF and NPAQ-Short not to reflect objectively measured PA adequately among the 54 included participants.

Study IV, the MIPAM RCT, only included 70 participants, and hence was not adequately powered. However, the intervention group walked on average 909 steps per day more than the control group and had a lower degree of loneliness, equivalent to a small to moderate effect size. Even though the finding for the primary outcome, daily steps, was not significant, the effect size can still be considered clinically relevant and the use of MI in combination with PAM-based interventions should be investigated further with adequately powered RCTs.

Results of study I, the effect of physical activity monitor interventions

A total of twenty-one studies (22 comparisons and 2,783 participants) were included (37–47,203–212), covering specific patient categories such as osteoarthritis (211), chronic obstructive pulmonary disease (42,44,205), cardiac patients (204,212), but also healthy or uncategorised populations (37–41,43,45–47,203,206–210). The median mean age in the studies was 70.5 years, and the median mean baseline step count, and thus the baseline activity level, was 5,268 per day. The median intervention length was 12 weeks with most (81%) studies providing daily feedback from PAMs (37–42,45–47,204,206–212).

Five studies were judged as having a low overall risk of bias (39,45,203,211,212), 10 studies were judged as raising some concerns (40,42–44,47,204,205,207,209,210), and six studies were judged as having a high risk (37,41,42,46,206,208). Judgments and support for judgment about each domain is presented in the full study report (175).

The overall SMD, number of study comparisons, number of participants and the level of heterogeneity can be seen for each outcome in number Figure 7 below.

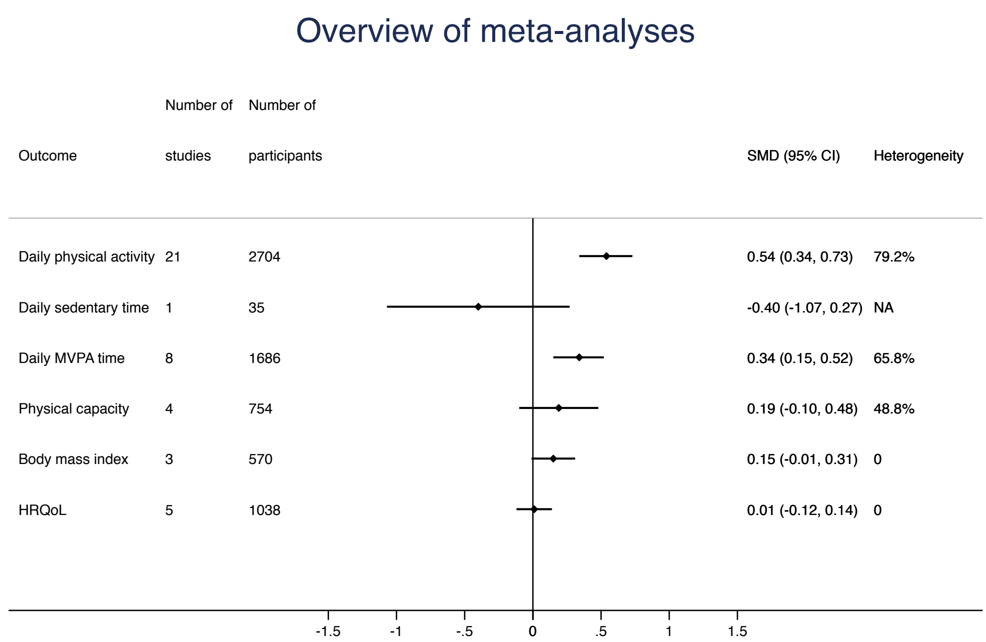


Figure 7. Overview of random effects meta-analyses for each outcome performed in **study I**. SMD: standardised mean difference, MVPA: Moderate to Vigorous Physical Activity, HRQoL: Health-related Quality of Life. Heterogeneity expressed as I^2 values. Positive SMD equals higher values in the intervention groups. Positive values are favoured for daily physical activity, MVPA, physical capacity, and HRQoL. Negative values are favoured for daily sedentary time.

The overall relative risk of adverse events was 0.91, (95% CI: 0.66; 1.25), I^2=0.0% p=0.942, with more adverse events in the control groups. Subgroup analyses on categorical variables (e.g. diagnoses, feedback frequency and risk of bias) are listed in the original study report for all outcomes (175). No significant differences in the subgroup analyses were observed with any outcomes. Sensitivity analyses (univariate meta-regressions) on continuous variables (e.g. age in years, sex distribution, percent of participants with walking aids, intervention length in weeks, or baseline PA) were performed, but none of these variables were significantly correlated with the effect size for any outcomes nor did they reduce the Tau^2 statistic. Egger's test revealed significant small study bias for the effect on daily PA, p=0.036, indicating overestimation of the SMD. The bias-adjusted result revealed a SMD of 0.37, (95% CI: 0.15; 0.59) down from 0.54, (95%CI: 0.34; 0.73) after filling the analysis with three fictive studies. No other small study bias was found for the other outcomes. The 11 study comparisons (1,219 participants) with passive control interventions had a significantly larger SMD (0.86, (95%CI: 0.53; 1.20)) compared; the 10 study comparisons (1,485 participants) with an active control intervention (0.22, (95%CI: 0.03; 0.41)).

The summary of findings for all outcomes is found below in Table 2, including certainty of the evidence for each outcome and translations from SMDs to mean differences.

Table 2. Summary of findings table from **study I**.

Physical activity monitor interventions compared to control interventions in older adults

Outcomes	**Anticipated absolute effects*** (95% CI)	**Standardised mean difference** (95% CI)	**№ of studies** (participants)	**Certainty of the evidence** (GRADE)	**Comments**
Physical activity - objectively measured or self-reported	The translated weighted mean difference was 1,297 steps per day (95% CI: 817; 1,753) with more steps in the intervention groups	SMD 0.53 (0.34; 0.73) in favour of the intervention	20 studies (21 study comparisons and 2,704 participants)	⊕⊕◯◯[1,2] Low	None
Time spent sedentary	The mean difference of weekly sedentary time was 44 minutes (37.1; 50.9) with the control group being more sedentary.	SMD 0.40 (-0.27; 1.07) in favour of the intervention	One study with 39 participants	⊕◯◯◯[4] Very low	None
Moderate to vigorous physical activity - objectively measured or self-reported	The translated weighted mean difference was hence 5.5 minutes per day (95% CI: 2.4; 8.4) with more MVPA in the intervention groups	SMD 0.34 (0.15; 0.52) in favour of the intervention	8 studies (1,686 participants)	⊕⊕⊕◯[1] Moderate	None
Physical capacity measured with walking tests	The translated weighted mean difference was 15 meters (95% CI: -8; 38) with more meters walked on a 6MWT in the intervention groups.	SMD 0.19 (-0.10; 0.48) in favour of the intervention	4 studies (754 participants)	⊕⊕◯◯[1,3] Low	None
Body mass index	The translated weighted mean difference was 0.72 kg/m^2 (95% CI: -0.048; 1.5) with the control group having the lowest BMI.	SMD 0.15 (-0.01; 0.31) in favour of the control intervention	3 studies (570 participants)	⊕⊕⊕◯[1] Moderate	None
Self-reported HRQoL – assessed with questionnaires	No mean value available as the studies used different outcome measures.	SMD 0.01 (-0.12; 0.14) in favour of the intervention	5 studies (1,038 participants)	⊕⊕◯◯[1,3] Low	None

Table 2. Summary of findings table from **study I**.

Physical activity monitor interventions compared to control interventions in older adults						
Outcomes	**Anticipated absolute effects*** (95% CI)		**Standardised mean difference** (95% CI)	**№ of studies** (participants)	**Certainty of the evidence** (GRADE)	**Comments**
Meeting the study specific recommended level of physical activity	No data		No data	No data	No data	No data
Adverse events and withdrawals	Calculated with 64 withdrawals of 881 participants in control groups	Calculated with 63 withdrawals of 1,149 participants in intervention groups	Relative risk for withdrawal due to illness or adverse events 0.91 (0.66; 1.25) with higher risk in control group	11 studies (1,927 participants)	⊕⊕⊕○[3] Moderate	None

*The absolute effects are calculated from the standardised mean differences and a relevant standard deviation according to the method section.
Abbreviations: HRQoL: Health related quality of life CI**:** Confidence interval; SMD**:** Standardised mean difference.
[1]: Downgraded by one level due to inconsistency (unexplained heterogeneity)
[2]: Downgraded by one level due to publication bias
[3]: Downgraded by one level due to imprecision of the results
[4]: Sparse data

GRADE Working Group grades of evidence
High certainty: We are very confident that the true effect lies close to that of the estimate of the effect. **Moderate certainty:** 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 certainty:** Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. **Very low certainty:** We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

Results of study II and study III, criterion validity of consumer available physical activity monitors and concurrent validity of physical activity questionnaires

For interpretation purposes, results from **study II** and **study III** are reported together.

Participants

In **study II**, 103 older adults volunteered to participate. Thirty-three percent of the participants were male, and they were on average 81 years of age, had a BMI of 28 kg/m^2 and 35 percent used a rollator to walk. They walked 255 meters in the six-minute test on average, equivalent to a self-paced speed of 2.6 km/h. For a full detailed list of characteristics, see the full study report (135). In **Study III,** 67 participants were considered eligible for inclusion from **study IV**. Of these, 13 were excluded after having insufficient number of days wearing the devices, leaving 54 participants for inclusion in the study. Forty-four percent of the participants were male, and they were on average 71 years of age, had a BMI of 27 kg/m^2 and one participant used a cane to walk, while no participants used rollators. On average, the participants reported their total PA to consist of 58% walking. Table 3 reports participants' number of daily steps, as well as IPAQ-SF and NPAQ-short results.

Table 3. Modified table from **study III**. Objectively measured steps and instrument scores.

Variable	**Total sample (n=54)**		
	Mean (SD)	**Median [IQR]**	
Daily steps	5,782.4 (3,005.4)	5,682.9 [3,422.4; 8,036.1]	
International Physical Activity Short Form Questionnaire	**Mean (SD)**	**Median [IQR]**	**Frequency of zero activity; n (%)**
IPAQ-SF Vigorous Intensity Minutes/Week	144.7 (390.9)	0.0 (0.0; 141.3)	30 (56%)
IPAQ-SF Moderate Intensity Minutes/Week	185.4 (418.0)	25.0 [0.0; 232.5]	25 (46%)
IPAQ-SF MET-Minutes/Week	2;972.1 (5;126.5)	1;410 [594.0; 3;318.8]	6 (11%)
IPAQ-SF MVPA time/Week	330.0 (781.1)	62.5 [0.0; 350.0]	18 (33%)
IPAQ-SF Sedentary Minutes/Day	329.4 (184.0)	300.0 [192.5; 465.0]	2 (4%)
IPAQ-SF Walking Minutes/Week	325.3 (465.0)	180.0 [0.00; 420.0]	20 (37%)
IPAQ-SF Physical Activity Categories			
Low activity level, n (%)		27 (50.0%)	
Moderate activity level, n (%)		11 (20.4%)	
High activity level, n (%)		16 (29.6%)	
Nordic Physical Activity Short Questionnaire	**Mean (SD)**	**Median [IQR]**	**Frequency of zero activity, n (%)**
NPAQ-short Vigorous Physical Activity Minutes/Week	142.0 (183.7)	79.5 [0.00; 172.5]	18 (33%)
NPAQ- short Moderate Physical Activity Minutes/Week	269.7 (399.7)	120.0 [0.0; 286.8]	15 (28%)
NPAQ-short MVPA Time Minutes/Week	411.6 (447.4)	240.0 [122.5; 592.5]	6 (11%)
NPAQ-short Physical Activity Categories			
Inactive		6 (11.1%)	
Not sufficiently physically active		7 (13.0%)	
Sufficiently physically active		10 (18.5%)	
Optimally physically active		31 (57.4%)	
NPAQ-short Compliance with WHO recommendations			
No		13 (24.1%)	
Yes		41 (75.9%)	

Abbreviations: SD: Standard Deviation, IQR: Inter Quartile Range, IPAQ: International Physical Activity Questionnaire, NPAQ: Nordic Physical Activity Questionnaire. MET: Metabolic Equivalent of Task, MVPA: Moderate to Vigorous Physical Activity, WHO: World Health Organization. Daily steps per day and height were found to be normally distributed. All other continuous variables were judged as having a non-normal distribution. Number of "zero" responses is defined as participants who reported no activity in a specific category. IPAQ-SF Categories: 'Low activity level' refers to participants who did not meet the criteria for moderate and vigorous intensity categories. 'Moderate activity level' refers participants who had a) three or more days of vigorous-intensity activity of at least 20 minutes per day, b) five or more days of moderate-intensity activity and/or walking of at least 30 minutes per day or c) 5 or more days of any combination of walking, moderate-intensity or vigorous intensity activities achieving a minimum total physical activity of at least 600 MET-minutes/week. 'High activity level' refers to participants who had a) vigorous-intensity activity on at least 3 days (20min minimum, achieving a minimum Total physical activity of at least 1500 MET-minutes/week or b) 7 or more days of any combination of walking, moderate-intensity or vigorous-intensity activities achieving a minimum total physical activity of at least 3000 MET-minutes/week. NPAQ-short Categories: Inactive participants were defined as participants with no moderate or vigorous physical activity, not sufficiently physically active (moderate physical activity/150 + vigorous physical activity/75)<1.0 and ((moderate physical activity /150 + vigorous physical activity /75)>0), sufficiently physically active ((moderate physical activity /150 + vigorous physical activity /75) ≥ 1.0 and (moderate physical activity /150 + vigorous physical activity /75)<2.0), optimally physically active ((moderate physical activity /150 + vigorous physical activity /75) ≥ 2.0). Compliance with WHO recommendations included sufficiently and optimally physically active participants.

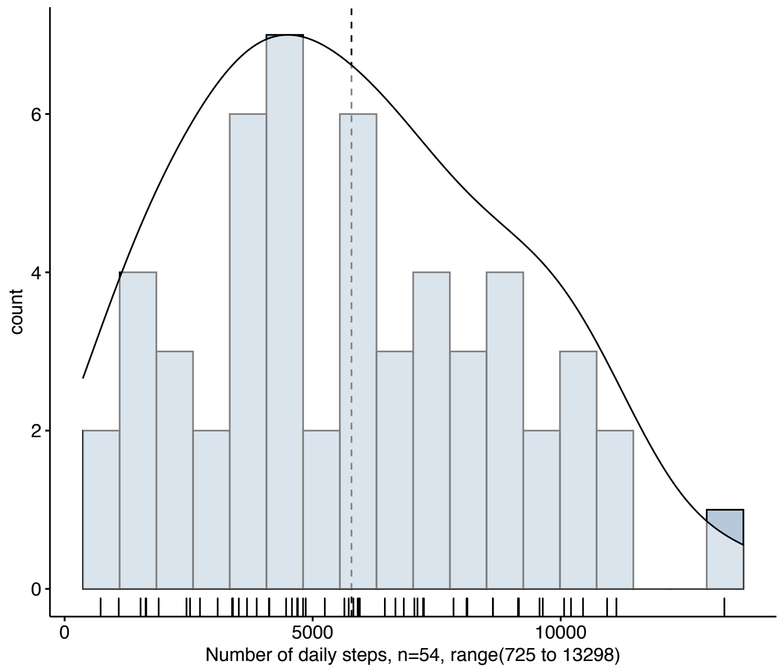


Figure 8. Histogram with density curve of the average step count from the 54 participants included in **study III**. The mean (standard deviation) step count was 5,782.4 (3,005.4) and the values range from 725 to 13,298.

Criterion validity of physical activity monitors

In **study II,** four consumer-grade PAMs currently on the market were identified as meeting the inclusion criteria: The Misfit Shine, Nokia GO, Jawbone UP Move and Garmin Vivofit 3. It proved impossible perform synchronisation between participants, and thus to extract data, for the Nokia Go, for which reason these PAMs were excluded from the study. Figure 9 below illustrates the percentage of excluded observations from the total available observations. Wrist measures were more likely to be excluded, p<0.001. There were significant differences in the number of deleted observations between the Garmin, Jawbone and Misfit PAMs, p<0.001. No between-group differences were found when comparing left and right measures, p=0.816.

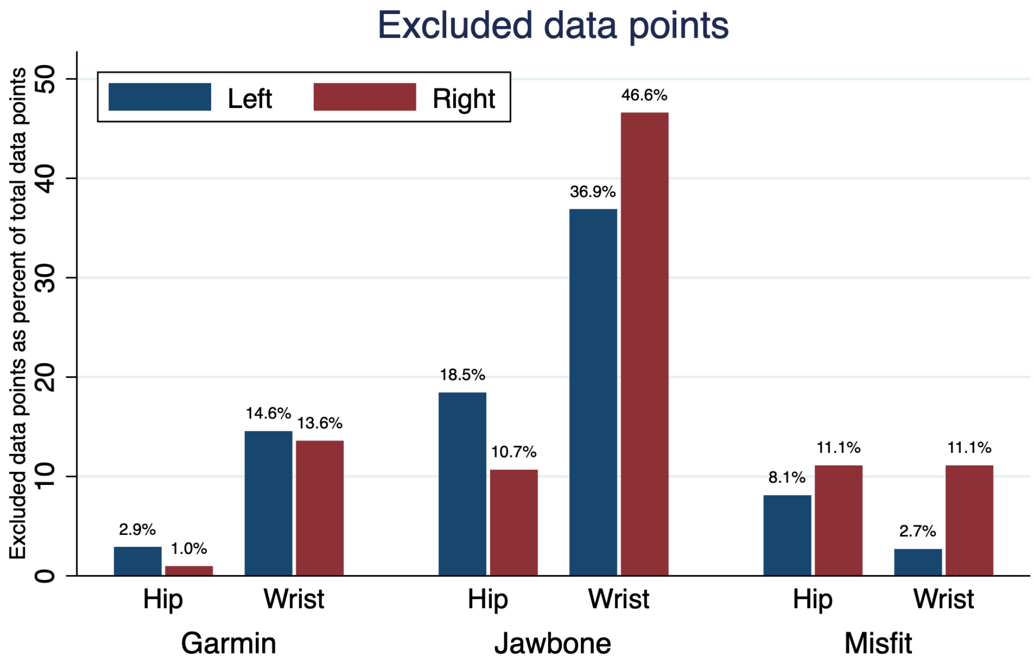


Figure 9. Bar chart of excluded data points as percentage of total available data points for the Garmin Vivofit 3 (n=103), Jawbone UP Move (n=103) and the Misfit shine (n=37 for left-worn and n=99 for right-worn) included in **study II**. The observations are divided into hip/wrist and left/right categories. This figure is similar to the figure in **study II**.

Comparing counts between the two testers found a good degree of interrater reliability (ICC(2,1) at 0.88 (95% CI 0.83; 0.92), with a mean difference of 4.42 steps (95% CI : 6.10; 14.91)). Detailed results on criterion validity ICC(2,1), mean difference and percentage measurement error for all PAMs on all positions can be found in the full study report. For hip-worn PAMs, 10 out of 18 possible combinations of brand, position and use of a rollator fulfilled the a priori hypothesis (of moderate criterion validity for all participants, good validity for participants without a rollator and poor validity for participants with rollators). For wrist-worn monitors, only one combination fulfilled the a priori hypothesis. The hip-worn Misfit Shine fulfilled the a priori hypothesis in four out of six combinations. The hip-worn Garmin Vivofit 3 fulfilled the a priori hypothesis in five out of six combinations. The hip-worn Jawbone UP Move fulfilled the a priori hypothesis in one out of six combinations. For the wrist-worn PAMs, no combination fulfilled the a priori hypothesis for criterion validity, except the right-worn Garmin Vivofit 3 for participants with rollators.

Figure 10 below illustrates the relationship between measurement error and the criterion (counted steps). The figures show negative slopes for all PAMs in participants without rollators and for hip-worn monitors for participants with rollators. For wrist-worn monitors in participants with rollators, the slope is horizontal and has larger measurement error, compared to the hip-worn PAMs.

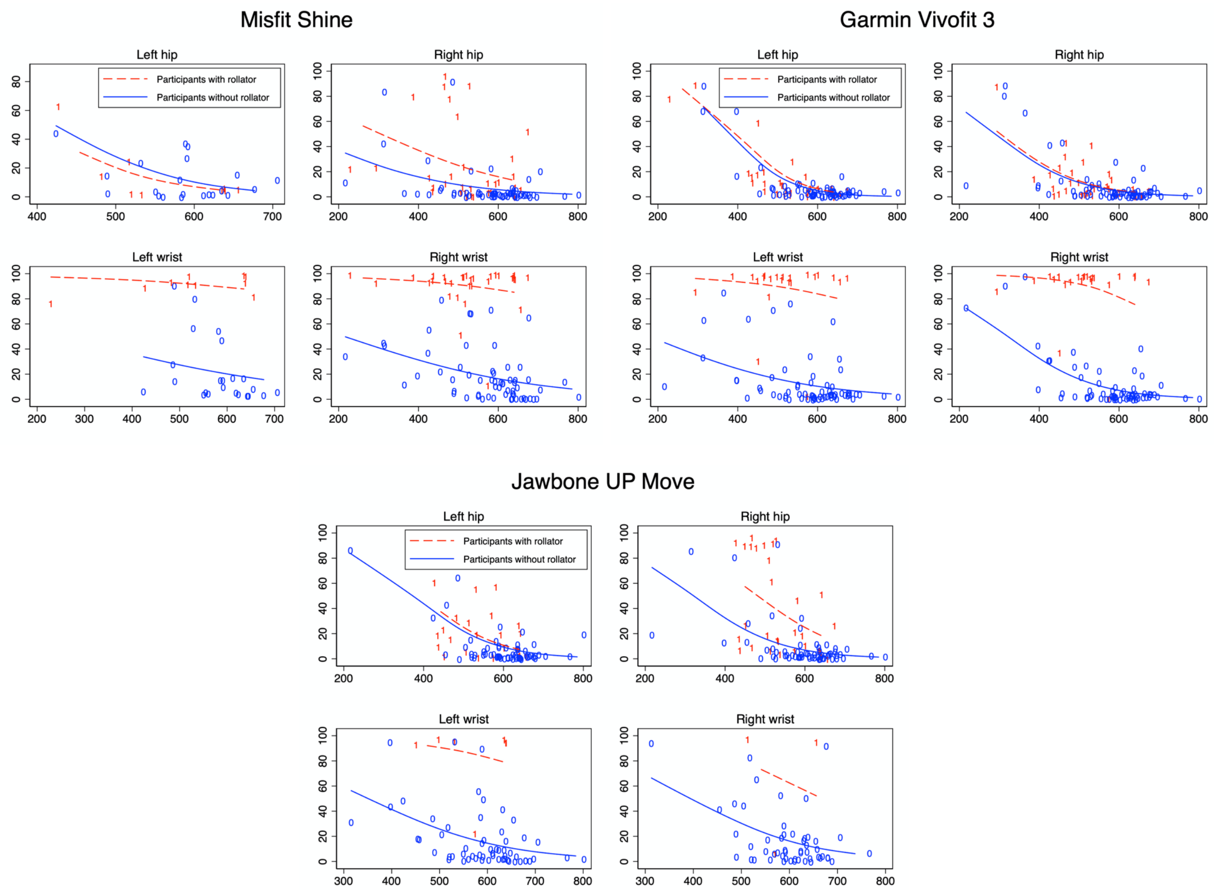


Figure 10 (a, b and c). Two-way scatter plots with logit link generalised linear models between absolute measurement error in % and observed steps of Misfit Shine (a), Garmin Vivofit 3 (b) and Jawbone UP Move (c) physical activity monitors. Each includes results from the left hip, right hip, left wrist and right wrist. Red digits "1" and lines equal participants with rollators and blue digits "0" and lines equal participants without rollators. Y-axis represents absolute measurement error in percentage as a response to the x-axis, which represents the number of observed steps. These figures are similar to the figures in **study II**.

Concurrent validity of physical activity questionnaires

Study III found that 'moderate intensity minutes', 'MVPA' and 'sedentary time' in IPAQ-SF had little or no correlation with daily steps (ρ<0.25). For the NPAQ-short, scores with little or no correlation included 'vigorous activity', 'moderate activity', 'MVPA' and 'compliance with WHO recommendations'. IPAQ-SF scores with fair correlation (0.25<ρ>0.5) include 'vigorous activity', and 'walking'. IPAQ-SF 'MET-minutes' showed moderate to good correlation (0.5<ρ>0.75) with daily steps. Figure 11 reports the concurrent validity between daily steps and the PAQ scores.

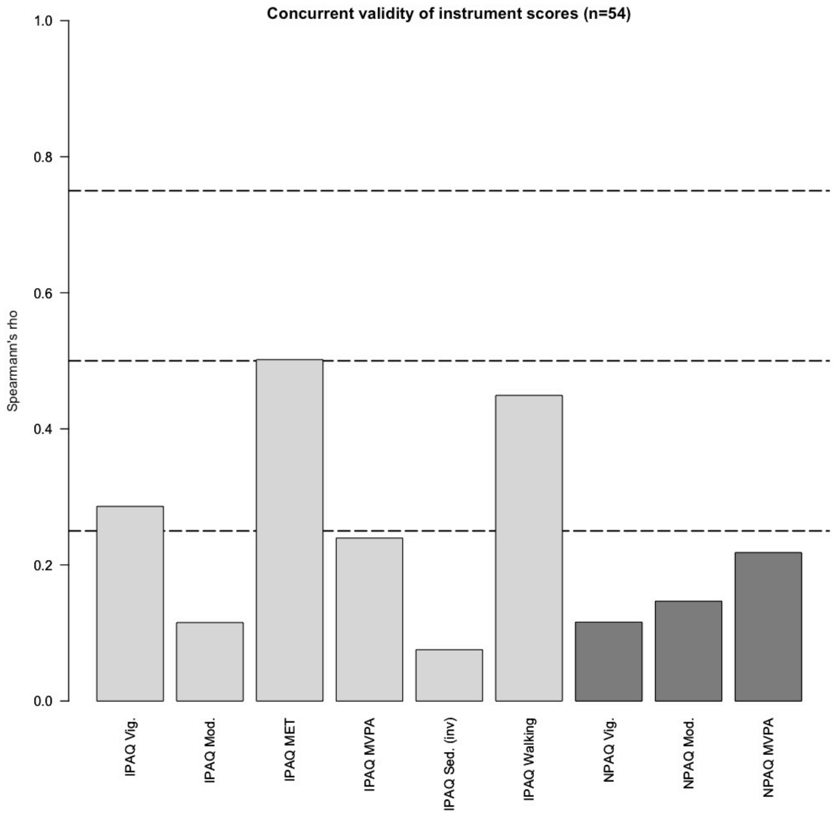


Figure 11. Performance of instrument scores expressed as the correlation between the instrument and objectively measured steps per day in **study III**. Light grey coloured bars represent IPAQ-SF constructs and dark grey coloured bars represent NPAQ-short constructs. Y-axis: Spearman's rank correlation coefficient (rho). Reference lines: below the lowest line indicates little or no correlation (ρ<0.25), above the lowest line indicates a fair correlation (ρ>0.25), above the middle line indicates a moderate to good correlation (ρ>0.5) and above the top line indicates a good to excellent correlation (ρ>0.75). IPAQ-SF Sedentary construct is inverse, meaning the true value is the negative equivalent of the illustrated. Abbreviations; IPAQ: International Physical Activity Questionnaire, Vig: Vigorous, Mod: Moderate, MET: Metabolic Equivalent of Task, MVPA: Moderate to Vigorous Physical Activity, Sed: Sedentary, NPAQ: Nordic Physical Activity Questionnaire. This figure is similar to the figure in **study III**.

As for the differences in daily steps between participants in the three IPAQ-SF categories, the study found that the mean step count in the 'low' category was 3,531 (95%CI: 1,565; 5,497) lower than the mean in the 'high' category. The mean in the 'low' category was 2,561 (95%CI: 332; 4,970) lower than the mean in the 'moderate category'. Finally, the mean in the 'moderate' category was 970 (95%CI: -1,471; 3,410) lower than the mean in the 'high' category. A one-way analysis of variance revealed that not all of the means were similar, p<0.001. No differences between the four NPAQ-Short categories were found, p=0.240.

An unpaired t-test revealed that the group adhering to WHO recommendations on average walked a further 1,628 steps per day (95%CI: 76; 3,180) compared to the non-adherent group. Distributions, means and 95% CIs are illustrated below in Figure 12.

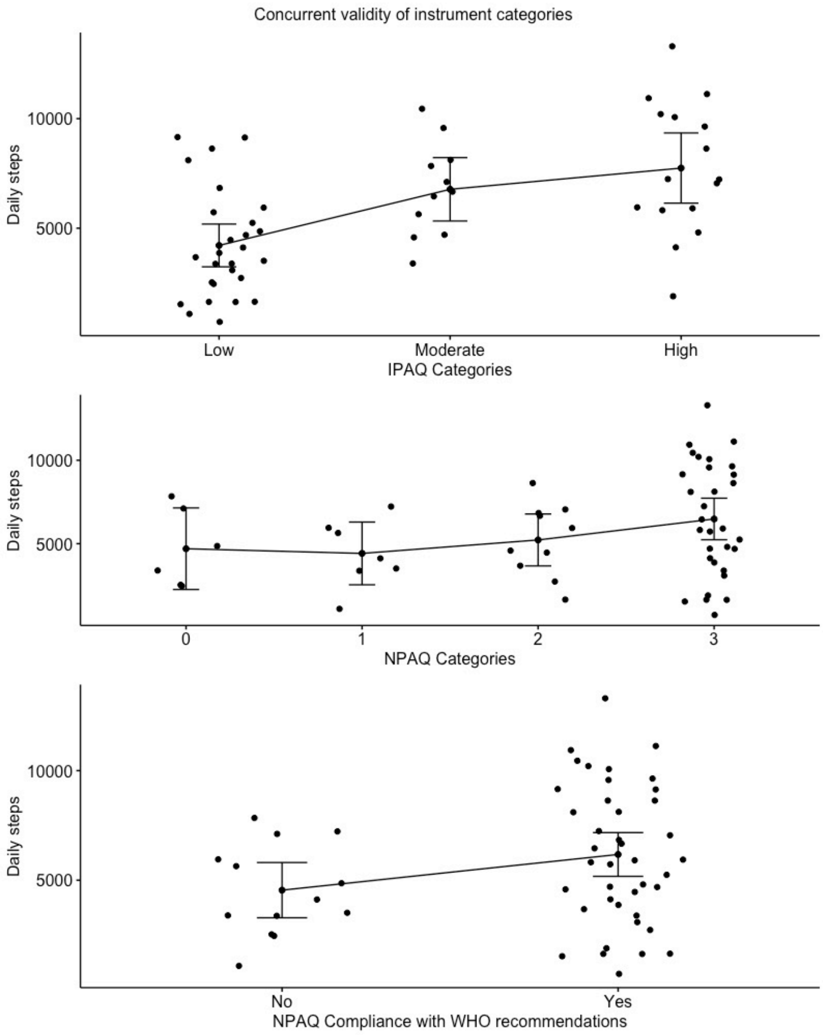


Figure 12. Means, confidence intervals and distributions of average steps per day of the International Physical Activity Questionnaire Short Form categorical construct and the Nordic Physical Activity Questionnaire Short categorical constructs. **NPAQ-short Categories**: 0) Inactive participants were defined as participants with no moderate or vigorous physical activity, 1) not sufficiently physically active (moderate physical activity/150 + vigorous physical activity/75)<1.0 and ((moderate physical activity /150 + vigorous physical activity /75)>0), 2) sufficiently physically active ((moderate physical activity /150 + vigorous physical activity /75) ≥ 1.0 and (moderate physical activity /150 + vigorous physical activity /75)<2.0), 3) optimally physically active ((moderate physical activity /150 + vigorous physical activity /75) ≥ 2.0). Compliance with WHO recommendations included sufficiently and optimally physically active participants. This figure is similar to the figure in **study III**.

Results of study IV, motivational interviewing and physical activity monitoring to increase the daily level of physical activity

Out of 79 possible participants, nine refused to participate after reading the study material, and 70 were included and randomised. Thirty-eight were allocated to the PAM group and 32 to the PAM+MI group. In the PAM group, 34 participants completed the 12 weeks and four participants discontinued. In the PAM+MI group, 28 participants completed the 12 weeks and four discontinued. Due to slow rate of inclusion and insufficient funding, the inclusion of participants was terminated in January 2020 and the trial did not reach the desired sample size of 128 participants. The CONSORT flow diagram is reported below in Figure 13.

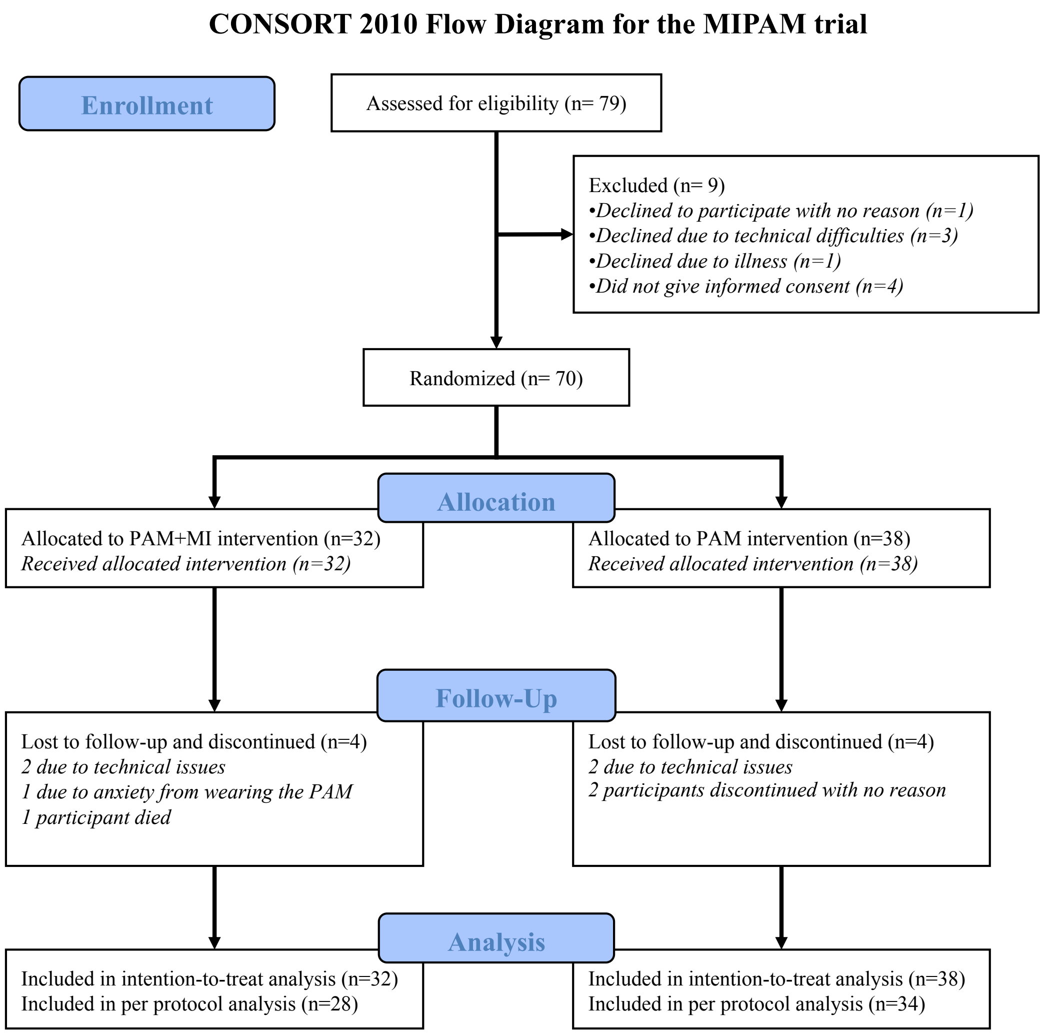


Figure 13. Consort flow diagram from **study IV**. This figure is similar to the figure in **study IV**.

Baseline data

Characteristics of the included participants are reported in full in Table 4.

Table 4. Modified table from **study IV**. Demographics and physical activity characteristics of included participants.

	Overall (n=70)	**PAM group (n=38)**	**PAM+MI group (n=32)**	***p***
Age, median [IQR]	72.0 [70.0, 74.0]	71.0 [70.0; 74.3]	73.0 [71.0, 74.0]	0.134
Sex, n male (%)	28 (40.0)	16 (42.1)	12 (37.5)	0.613
BMI, mean (SD)	27.2 (4.4)	27.3 (4.9)	27.1 (3.9)	0.581
In pain, n (%)	25 (37.3)	9 (25.0)	16 (51.6)	0.046
Wants to be more physically active, n (%)				0.259
Yes	56 (83.6)	28 (77.8)	28 (90.3)	
No	3 (4.5)	2 (5.6)	1 (3.2)	
Do not know	8 (11.9)	6 (16.7)	2 (6.5)	
Have used or uses physical activity monitor, n (%)	22 (32.8)	12 (33.3)	10 (32.3)	0.997
UCLA Loneliness Scale Sum, mean (SD)	32.9 (8.6)	33.47 (9.51)	32.27 (7.46)	0.399
EQ-5D-5L				
Problems with mobility, n (%)	27 (40.1)	13 (36.1)	14 (46.7)	0.373
Problems with self-care, n (%)	4 (6.1)	2 (5.6)	2 (6.7)	0.995
Problems with usual activities, n (%)	19 (28.8)	9 (25.0)	10 (33.3)	0.442
Problems with pain and discomfort, n (%)	43 (65.2)	20 (55.6)	23 (76.7)	0.087
Problems with anxiety and depression, n (%)	13 (19.7)	7 (19.4)	6 (20.0)	0.995
EQ Visual Analogue Scale, median [IQR]	80.0 [70.0; 90.0]	85.0 [70.0; 90.0]	80.0 [70.0; 90.0]	0.438
Outcome Expectancy for Exercise-2 Scale Sum, mean (SD)	51.6 (6.9)	50.3 (7.27)	53.1 (6.1)	0.074
Self-Efficacy for Exercise Scale Sum, mean (SD)	60.5 (19.8)	59.4 (20.15)	61.8 (20.0)	0.442
Baseline steps per day, mean (SD)	5881.1 (2948.9)	6029.4 (3009.58)	5705.2 (2913.0)	0.451
International Physical Activity Questionnaire Short Form				
Minutes of vigorous activity per day, median [IQR]	0.0 [0.0; 24.1]	0.0 [0.0; 19.3]	0.0 [0.0; 24.1]	0.581
Minutes of moderate activity per day, median [IQR]	0.0 [0.0; 24.1]	0.0 [0.0; 19.3]	0.0 [0.0; 24.1]	0.581
Minutes of MVPA per day, median [IQR]	0.0 [0.0; 48.2]	0.0 [0.0; 38.6]	0.0 [0.0; 48.2]	0.581
Minutes of walking activity per day, median [IQR]	27.9 [0.0; 327.9]	23.6 [0.0; 422.1]	31.1 [0.0; 214.3]	0.643
MET-minutes per day, median [IQR]	316.9 [63.3; 1386.0]	394.1 [40.9; 1432.0]	254.6 [102.5; 850.5]	0.570
Minutes of sedentary time per day, median [IQR]	303.0 [210.0; 480.0]	303.0 [240.0; 435.0]	316.0 [202.5; 480.8]	0.676
IPAQ physical activity categories, n (%)				0.690
High	18 (27.3)	10 (27.8)	8 (26.7)	
Low	26 (39.4)	14 (38.9)	12 (40.0)	
Moderate	22 (33.3)	12 (33.3)	10 (33.3)	
Nordic Physical Activity Questionnaire short				
Minutes of MVPA per day, median [IQR]	30.0 [12.9; 68.6]	38.6 [8.2; 82.5]	25.7 [17.5; 46.1]	0.227
Minutes of moderate activity per day, median [IQR]	11.4 [0.0; 24.6]	7.9 [0.0; 34.3]	12.1 [0.0; 19.8]	0.651
Minutes of vigorous activity per day, median [IQR]	17.1 [0.0; 34.3]	19.3 [0.0; 58.4]	13.6 [0.0; 21.4]	0.225
NPAQ physical activity categories, n (%)				0.231
Inactive	8 (12.1)	4 (11.1)	4 (13.3)	
Insufficiently physically active	10 (15.2)	7 (19.4)	3 (10.0)	
Sufficiently physically active	12 (18.2)	4 (11.1)	8 (26.7)	
Optimally physically active	36 (54.5)	21 (58.3)	15 (50.0)	
NPAQ compliance with WHO recommendations, n (%)	48 (72.7)	25 (69.4)	23 (76.7)	0.490

Abbreviations: BMI: Body Mass Index, PAM: Physical Activity Monitor, EQ-5D: EuroQol Research Foundation Five Domains, UCLA: University of California Los Angeles, OEE: Outcome Expectancy for Exercise, SEE: Self Efficacy for Exercise. IPAQ-SF: International Physical Activity Questionnaire-Short Form, NPAQ: Nordic Physical Activity Questionnaire-Short Form, MVPA: Moderate to Vigorous Physical Activity, SD: standard deviation, IQR: interquartile range, IPAQ-SF: International Physical Activity Questionnaire-Short Form, NPAQ: Nordic Physical Activity Questionnaire-Short Form, MVPA: Moderate to Vigorous Physical Activity. Test for between-group difference in normal distributed continuous variables (BMI, UCLA Loneliness Scale Sum Score, Outcome Expectancy for Exercise-2 Scale Sum Score, Self-Efficacy for Exercise Sum Score and Baseline Daily Steps) were performed with unpaired t-test, tests for between-group difference in non-normal distributed continuous (age, % of total activity from walking, EQ Visual Analogue Scale, all IPAQ and NPAQ scores) variables were performed with Mann-Whitney U test, test for between group difference in categorical or binary variables with Chi2 test, p-values comparing the PAM and the PAM+MI group ≤ 0.05 are considered significant.

The median days of missing PA data was 5 [IQR: 1; 29], with 6 [IQR: 1; 32] days in the PAM group and 4.5 [IQR: 0.75; 26] days in the PAM+MI group. No significant between-group difference was found, p=0.484. For all randomised participants, data for four were imputed for average daily steps. The IPAQ-SF 'MVPA' and 'sedentary time', NPAQ-Short 'MVPA', EQ-VAS, UCLA Loneliness Scale Sum Score, and SEE-DK Sum Score all had data for six participants imputed. For seven participants, data was imputed for IPAQ-Short minutes of walking per day and OEE2-DK Sum.

The results of the adjusted multiple regression models can be found in Table 5 below, and Figure 14 illustrates per protocol scores for the two study arms through the study period.

Table 5. End point scores and results from multiple regression models on outcomes from **study IV.**

	End point scores		*Adjusted between group difference from multiple regression model*		
Outcome	**PAM group (n=38)** mean (95%CI)	**PAM+MI group (n=32)** mean (95%CI)	**Coefficients (difference)**	**95%CI**	***p***
Average daily steps	5,837.7 (4,932.8; 6,742.6)	6,492.9 (5,472.5; 7,513.3)	909.0	(-71.0; 1889.0)	0.07
IPAQ-SF					
MVPA minutes per day	53.9 (15.3; 92.5)	34.4 (5.2; 63.6)	-0.2	(46.3; 45.8)	0.992
Minutes of walking per day	149.2 (59.1; 239.3)	218.5 (111.5; 325.5)	78.1	(-6.1; 217.3)	0.266
Minutes of sedentary time per day	358.5 (303.6; 413.4)	335.0 (273.0; 397.0)	-40.3	(-102.8; 22.1)	0.201
NPAQ-Short					
MVPA minutes per day	72.5 (41.0; 104.0)	66.6 (40.1; 93.1)	-3.8	(-45.3; 37.7)	0.856
EQ-VAS	80.6 (76.0; 85.1)	81.6 (78.2; 85.1)	2.9	(-1.9; 7.7)	0.227
UCLA Loneliness Scale Sum Score	32.8 (29.6; 36.0)	30.2 (27.4; 33.0)	-2.3	(-4.5; -1.24)	0.04
Self-Efficacy for Exercise Sum Score	52.5 (45.9; 59.1)	55.3 (45.9; 60.4)	3.5	(-4.3; 11.2)	0.375
Outcome Expectancy for Exercise-2 Sum Score	51.3 (48.5; 54.2)	53.2 (50.5; 56.0)	2.0	(-2.0; 6.0)	0.320

Abbreviations: IPAQ-SF: International Physical Activity Questionnaire Short Form, NPAQ-Short: Nordic Physical Activity Questionnaire Short, EQ-VAS: EuroQol Visual Analogue Scale, UCLA: University of California, Los Angeles. For four participants, data was imputed for average daily steps. Six participants had data imputed for IPAQ-SF MVPA and minutes of sedentary time per day, NPAQ-Short MVPA minutes per day, EQ-VAS, UCLA Loneliness Scale Sum Score, and SEE Sum Score. Seven participants had data imputed for IPAQ-Short minutes of walking per day and OEE-2 Sum. End point scores are unadjusted. The primary analysis is the multiple linear regression model adjusted for baseline score, baseline steps, and sex. Coefficients>0 means a higher value in the PAM+MI group. P-values<0.05 are considered significant.

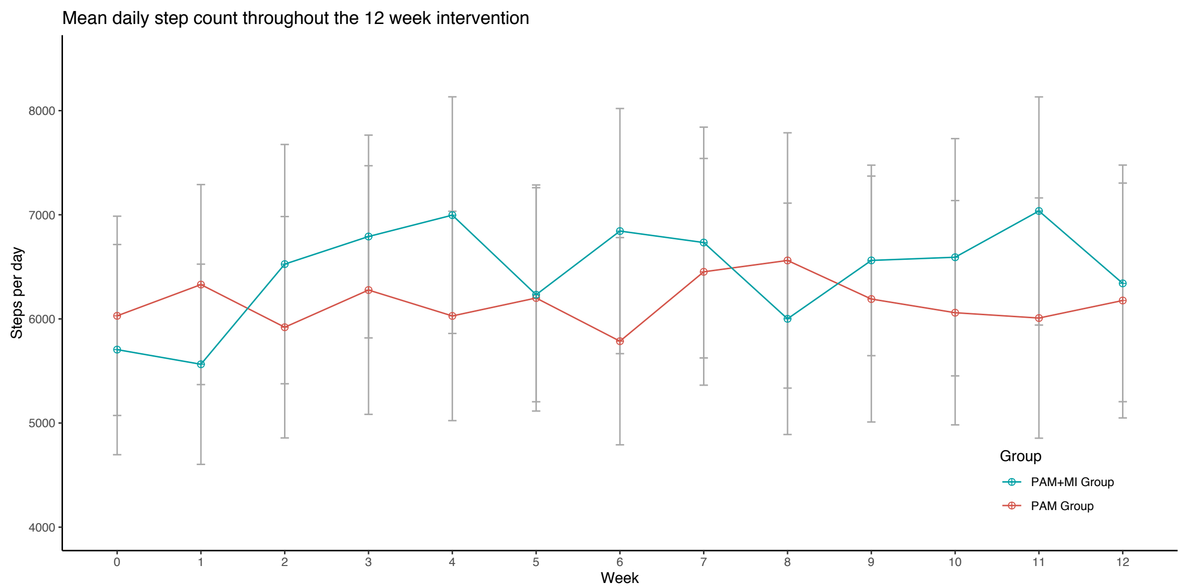


Figure 14. Unadjusted mean daily step counts throughout the 12-week intervention in **study IV**. Week 0: baseline week. Intervention period: week 1 to week 12. Circles represent mean values and error bars represent 95% confidence intervals. Unimputed data are used. This figure is similar to the figure in **study IV**.

Among the 28 PAM+MI complete case participants, 23 (82.1%) received all seven MI calls, four (14.3%) received six calls and one participant (3.6%) received only four calls. In total, 170 calls with an average length of 18.4 minutes were delivered to the PAM+MI group. Six MI calls were audiotaped and coded by the two coders using the Motivational Interviewing Treatment Integrity Scale. The median Cultivating Change Talk global score was 3.5, the median Softening Sustain talk score was 4, the median Partnership score was 4, and the median Empathy score was 4. The ratio of Reflections to Questions was 1.3.

The eight participants who discontinued the intervention differed significantly from the complete cases. They were older, at 78.5 years [IQR: 74.0; 81.5] versus 72.0 years [IQR: 70.0; 74.0], p=0.035, exclusively female (54.8% female in complete case versus 100% female in discontinued participants, p=0.038), and had a different use of walking aids (one rollator user and no cane users in the complete cases versus one cane user and no rollator users in discontinued participants, p=0.006). No other significant or clinically relevant differences were found for other baseline variables.

A post-hoc power calculation of the primary analysis revealed a power level of 24.6%. This analysis included 70 participants, an effect size of 909 steps, a SD of 2,948 steps, and an alpha level of 0.05.

Two participants, both allocated to the PAM+MI group (6.3%), discontinued due to adverse events (one died and one had increasing anxiety from wearing the PAM, which triggered existing mental illness). There was no significant between group-difference in the proportions of adverse events in the groups (0% in the PAM group versus 6.3% in the PAM+MI group, p=0.400).

Post-hoc sensitivity analyses and illustrations specifically conducted for the thesis

As previously described, the thesis includes post-hoc analyses to allow for a deeper discussion of the study results. The illustrations are not included in the published paper of **study IV**.

The likelihood of a participant having less than seven valid days of data in the intervention period and thus having their average step count imputed was investigated parametrically as well as non-parametrically. The participants with imputed data for PA were on average 8.5 years older (95% CI: 6.6; 10.7) than the participants who did not have their intervention step count imputed. Similarly, the number of missing days in the intervention period was found to be dependent on age, as the coefficient from a univariate linear model was 1.96 missing days per added year of age (95%CI: 0.56; 3.36). However, the normality of the distribution of both variables in the univariate linear model (age and number of missing days), as illustrated in Figure 17, can be questioned, which favours a non-parametric analysis over the parametric analyses. A Kendall's Tau coefficient was calculated to be 0.10, p=0.253.

The number of missing days was not dependent on the probability of reporting adverse events, illness or similar. The 56 participants who completed the post-intervention survey without reporting any illness or other, similar adverse events had a median amount of missing days of 5. The eight participants who completed the post-intervention survey and did report an adverse event, illness or similar, had a median amount of missing days of 2.5. A Wilcoxon test revealed no significant difference between the groups, p=0.362.

Between-group difference sensitivity analyses without multiple imputation revealed a between-group difference of 889 steps (95%CI: -99; 1877), p=0.077, for the complete case analysis and 825 steps (95%CI: -110; 1762), p=0.08, for the intention-to-treat analysis with last observation (baseline week) carried forward.

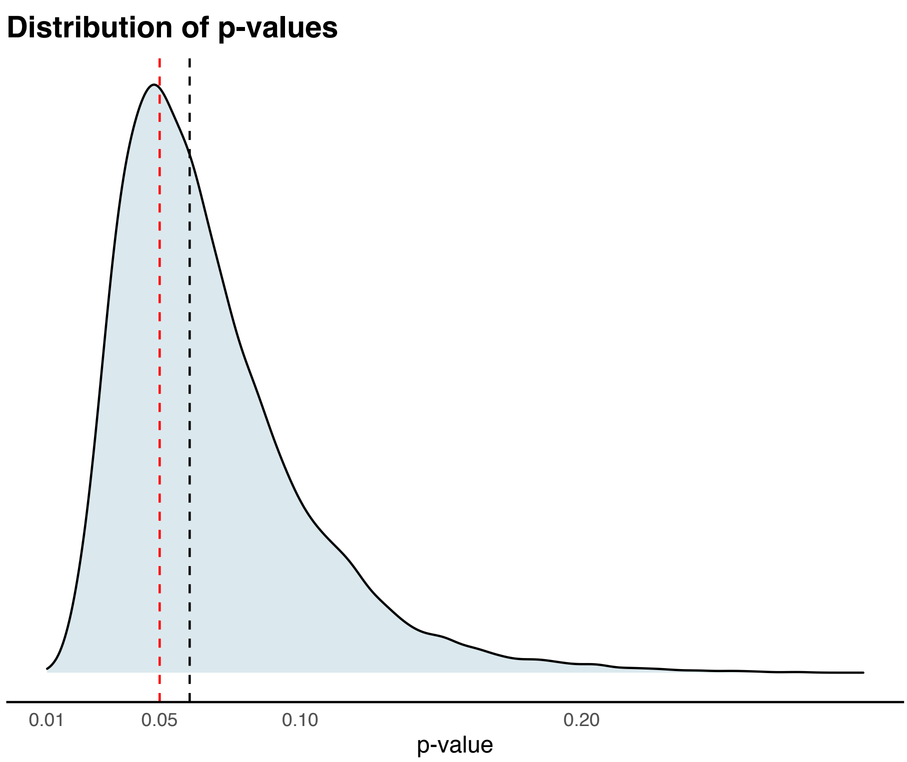


Figure 15. Density plot illustrating the distribution of p-values for between-group differences obtained from the multiple regression model in **study IV** with 5000 multiple imputations with 5 iterations for predictive mean matching. The red line equals the alpha level on 0.05. The black line equals the median p-value on 0.061.

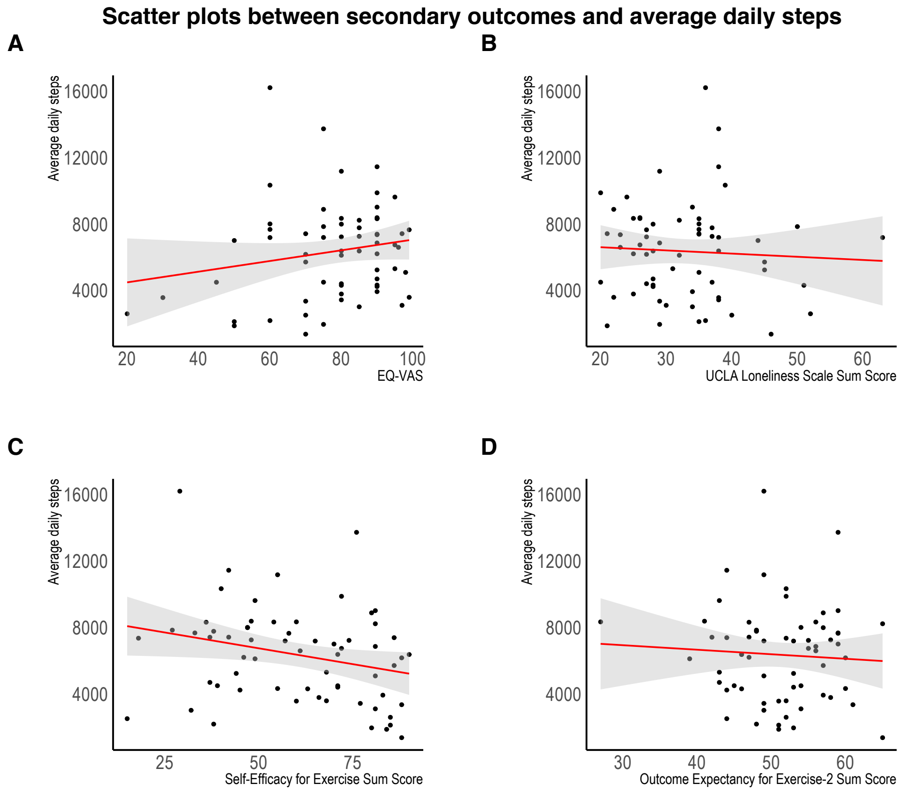


Figure 16. Scatter plots with best fitted lines and 95% confidence interval for average daily step count throughout the intervention period of **study IV** and A) EuroQol Visual Analogue Scale (EQ-VAS), b) University of California, Los Angeles (UCLA) Loneliness Scale Sum Score, c) Self-Efficacy for Exercise Sum Score, and D) Outcome-Expectancy for Exercise-2 Sum Score.

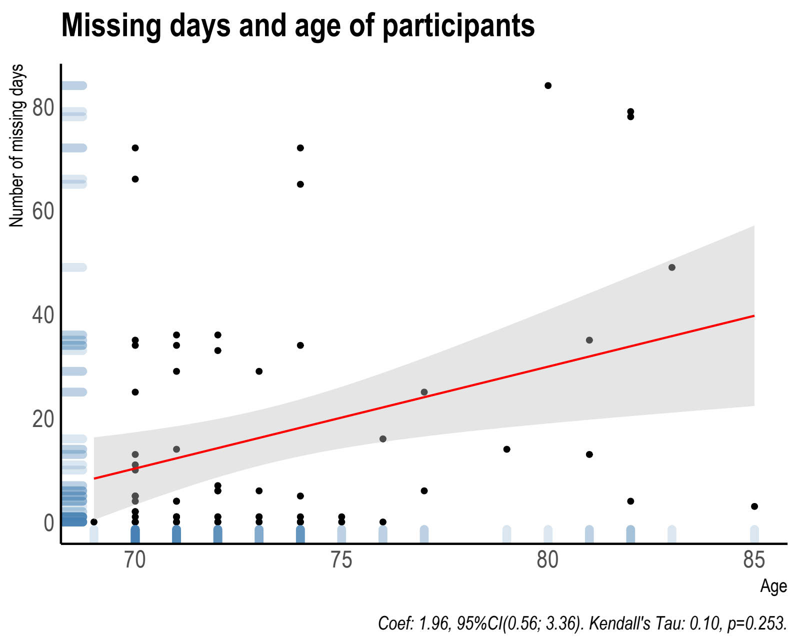


Figure 17. Scatter plot with best fitted line, 95% confidence interval, and rug margin plots of the distributions for number of missing days extracted from the physical activity monitor and age of the participants in **study IV**. The regression coefficient was 1.96, (95%CI: 0.56; 3.36) and the non-parametric equivalent (Kendall's Tau) was 0.10, p=0.253.

DISCUSSION

Summary of main findings

Four different studies were conducted in the course of this PhD project. The main results are presented in Figure 18, which illustrates the flow of the thesis. The systematic review and meta-analysis found a low quality of evidence for a moderate effect of PAM-based interventions on the daily PA level of older adults and a moderate quality of evidence for a small effect on daily MVPA time. The heterogeneity of results was not affected by anything apart from whether the comparisons were made between intervention groups and passive control groups, or active control groups. The criterion validity of consumer-grade PAMs in a population of older adults with and without rollators is mainly affected by the position, as the hip-worn PAMs were found to be superior to wrist-worn among the older adults in **Study II**. The hip-worn PAMs did not differ in terms of measurement error or validity, whereas the wrist-worn monitors were found to be associated with inadequate validity and a high degree of measurement error. The concurrent validity of the IPAQ-SF and NPAQ-Short was found to be inadequate, as the PAQs did not reflect objectively measured PA, and hence their use cannot be supported among older adults. When adding MI to a PAM-based PA intervention among community-dwelling older adults, MI was found to be effective, however not significantly so, in increasing the daily step count by 909 steps and decreasing the degree of loneliness equivalent to a small to moderate effect.

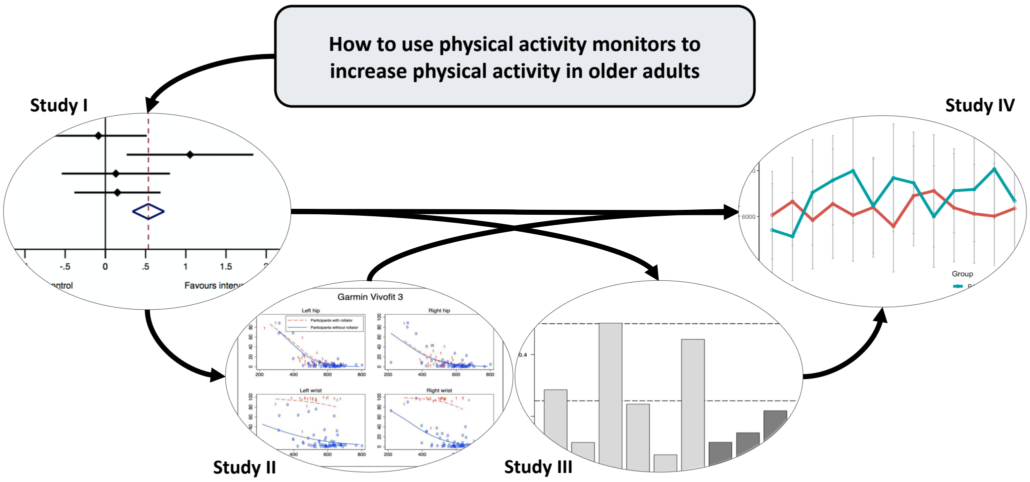


Figure 18. Illustration of the primary results of the four studies. **Study I** found a moderate effect of physical activity monitoring in general, **study II** found adequate criterion validity of hip-worn consumer-available physical activity monitors, **study III** found inadequate concurrent validity of the International Physical Activity Questionnaire Short Form and the Nordic Physical Activity Questionnaire Short among older adults, and **study IV** found a non-significant but clinically relevant effect of adding motivational interviewing to a physical activity monitor based intervention. Arrows illustrate how the results from **study I** was used in the design of the other studies and the results from **study II** and **study III** was used in the design and interpretation of **study IV**.

Discussion of main results

In this section the results and findings of the four studies are discussed with what is already known about this area of research. This section includes a discussion of the inadequate validity of PAQs among Danish older adults, the validity of consumer-available PAMs, the effectiveness of PAMs among older adults in general, and finally how motivational interviewing can be used in combination with PAM-based interventions.

Self-reported physical activity questionnaires among older adults

As mentioned in the introduction, validity is often inversely associated with applicability in objective instruments. PAQs are thus often used as time- and cost-efficient alternative, even though they are limited by recall and social desirability bias (84–87). The correlation coefficients between self-reported PA and objectively measured PA found in **study III** were lower than what has previously been reported in studies of older adults (90,93) while the validity of specifically the walking domain of the IPAQ-SF was found to be somewhat similar to a validation study reporting a strong correlation between the domain and daily steps (94). A possible explanation can lie in the use of objectively measured steps as the criterion, whereas two of the studies, those with higher correlation coefficients in most of the domains, used accelerometer counts or MVPA as their criterion, which might limit the comparison (90,213). Nevertheless, even though the walking domain of IPAQ-SF had a comparable level of validity with another study, 37% percent of the participants in **study III** reported zero walking activity. This highlights some of the issues with the instrument. No participants walked fewer than 725 steps per day, which is simply illustrated in the histogram with the density curve in Figure 8. Another explanation of lower correlations found in **study III**, could be that the other studies used younger participants (mean age of 65, 67 and 43 years respectively, compared with 71 years in **study III**) (214,90,93). However, despite what other studies might have found, the concurrent validity of the IPAQ-SF and the NPAQ-Short was inadequate and the number of zero responses was high. This suggests a problem with the ease of comprehension and face validity of the PAQs with this study population, and hence also the study population from **study IV**. In conclusion, the results of the IPAQ-SF and NPAQ-short, in the form of single domains and categories as used in **study III** and **study IV,** are not a suitable basis to draw conclusions about the PA level of the participants or even to distinguish between categories of activity levels among the participants. Future studies should validate the results of the questionnaires, rather than trusting self-reported PA alone. Furthermore, it may useful to include thorough guidance on how to answer the questions with more detailed descriptions of the constructs or to adapt version of the questionnaires for the elderly and thus to accept that self-reported PA among older adults should not be compared directly to results from younger populations (215).

Consumer-available physical activity monitors as outcome measurements

PAMs that are used as outcome measurements in clinical trials should obviously be subject to the same standards as any other medical outcome instrument. As described in the background section, the COSMIN and COMET collaborations are initiatives to help clinicians and researchers navigate and choose the right outcome measurement instruments (142,172). As also described in the background, the criterion validity (whether the results reflect the gold standard measurement) is especially important when using consumer-available PAMs as outcome instruments (142,216). **Study II** evaluated four such consumer-available PAMs, as they had been found relevant to use among older adults. As reported in **study II**, the Nokia GO was excluded due to feasibility issues and the Jawbone UP Move, Misfit Shine and Garmin Vivofit 3 did not differ significantly in measurement error or criterion validity when worn on the hip. The criterion of choice was visually counted steps, which distinguishes this study from others that have tended to use research-grade accelerometers as their criterion measures (71–77). For free-living conditions and estimating MVPA, research-grade accelerometers should continue to be the gold standard and criterion measure. However, for step counting, visually counted steps should be considered the criterion of choice in validation studies as research-grade accelerometers, like consumer-available PAMS, rely on the software and step-detection algorithms. One study with older adults even reported a consumer-available PAM to outperform a research-grade PAM in step detection (73). In addition to that, it should be noted that one can never be completely certain on the criterion measure. **Study II** chose the method of using two testers counting independently to limit measurement error. Even though even more careful and prodigious methods (e.g. video validation of all steps) could have been applied and would conceivably have increased the precision, the criterion measure arguably proved reliable because of its good interrater reliability, as reported in **study II.**

Study II also reported that the wrist-worn PAMs did not adequately measure the number of steps taken in populations including rollator users, which is illustrated in the logit link model from Figure 10. A nearly 100% measurement error in rollator users indicates that the lack of arm movement when using a rollator prevents the wrist-worn PAMs from detecting the steps. Furthermore, the negative slopes in the logit link models for all devices indicate that gait speed is also associated with the validity of the devices, which is in line with previous studies that have reported a greater amount of measurement error in slower walking participants (77,78). A validation study by Thorup et al. even reported a minimum walking speed of 3.6 km/hour for a consumer-available PAM to adequately measure steps among adults. Even though the analyses from **study II** were not used to determine a specific lower limit for walking speed, researchers should carefully consider the expected gait speed and use of assistive devices before choosing a specific consumer-available PAM as outcome instrument and not assume that all PAMs or placements will be equally valid with all populations.

Effectiveness of physical activity monitors as the facilitator for increasing physical activity levels

Beyond their use as outcome measurements, PAMs also have the ability to facilitate and motivate behavioural change, as per the aim of the thesis. **Study I** was the first systematic review to pool the results of the available evidence investigating the effect of PAM-based interventions among older adults. The main finding of this systematic review was the low quality of evidence for a moderate effect equivalent to the intervention groups on average walked 1,297 steps more daily. This positive result is in line with the JAMA systematic review by Bravata et al. from 2007 that estimated the effect size to be 2,491 steps per day in a population 20 years younger than the one included in **study I** (36). Currently, no better or updated systematic review of Bravata et al. exists. As it should be expected that an innovative and emerging field will change in 13 years, large scale systematic reviews investigating the effect of PAM-based interventions among adults are encouraged (217).

In **study I** it was not possible to explain any heterogeneity with pre-specified sub-group or sensitivity analyses. This resulted in a downgrading of the quality of evidence, but also meant that **study I** did not find any evidence to suggest that specific sub-populations experience fewer benefits than others. Hence, the recommendation about using PAMs for older adults to increase the level of PA applies to all older adults. Nevertheless, the studies with passive comparators had a significantly larger effect size compared to the studies with active comparators, indicating that PAM-based interventions are indeed effective, but mainly when compared to no interventions. Passive comparators include, but are not limited to, no-treatment or waitlist groups and should only be used to investigate whether the intervention works at all (218). Notably, when considering how the effect size of a study is calculated, the results from the control groups contribute to the overall result of a meta-analysis equally as much as the results from the intervention group. It is important to perform this type of sub-group analysis to obtain a transparent estimate of the effect of the PAM-based interventions and to avoid overestimated expectations of absolute effects (219). Nonetheless, the moderate effect size, equivalent to approximately 1,300 daily steps among older adults, is highly relevant as modest increments might have a substantial relevance when differences of 1,000 steps per day have been reported to lower the risk of all-cause mortality with as much as 11% (22).

The feasibility and effectiveness of motivational interviewing in combination with physical activity monitors

Study IV, the MIPAM trial, was the first study to investigate the effect of MI as an add-on intervention to a PAM-based, PA-promoting intervention among older adults. The main limitation of this study was the insufficient study sample and statistical power due to a lack of funding and a slow inclusion rate. Nevertheless, the PAM+MI group had a non-significant larger daily step count compared to the PAM group (909 steps, (95%CI: -71; 1889)). As is evident from the confidence interval, this suggests a high probability that MI is effective in increasing the daily PA level when added to a PAM-based intervention. This is in line with the findings of other randomised controlled trials investigating how MI can be used to increase PA levels (64,65) and the feasibility of MI in combination with PAMs (66). Unfortunately, not many high quality RCTs have been published with this focus, and as our general population of older adults differs from the populations of older patients (64,65,148), no conclusions about actual effect sizes can be drawn, which was also concluded in a recently published systematic review (220). MI among older adults appears efficacious in other research areas (smoking cessation, glycaemic control, blood pressure, cholesterol, and diet) and it would follow that MI can be used with good effect in older adults (55), similar to younger populations (221). However, naturally, with behavioural change interventions, heterogeneity exists among the individual study effects (55,148,220), and one size does not fit all, as this study made apparent. **Study IV** was designed as an add-on superiority study investigating the effect on PA. MI is a highly complex intervention where factors such as therapist skills, client change talk, personal- and relational factors might affect the delivery and effect of the intervention *(222)*. The literature on PA behaviour change among older adults is lacking in MI studies, and at the moment, evidence from other research areas need to be used to evaluate the results. However, as only 32 participants were included in the intervention group, the results of **study IV** do not have the statistical power to allow for post-hoc analyses on MI call dose, stages of change of individual participants or personal factors, such as baseline levels of self-efficacy or outcome expectancy. Consequently, the overall effect and between-group difference should be used to inform future high quality RCTs, and within-group participant level behaviours should not be extrapolated.

Previously studies have found associations between self-efficacy (223–225), outcome-expectancy (223,226), loneliness (227), and HRQoL (228,229) with PA levels among older adults. As illustrated in the scatterplots in Figure 16, none of the above-mentioned variables were linearly associated with the PA level in the intervention period among the participants of **study IV**. Some associations had been expected in planning the study and selecting the secondary outcome measurements, which is partly why MI was expected to increase the effects of the PAMs. Furthermore, no between-group differences in secondary outcomes were found after the intervention apart from the level of loneliness, which was found to be 2.3 (95%CI: 1.24; 4.5) points lower in the PAM+MI group, equivalent to a small to moderate effect size (24). This could have been explained by the increased PA levels in the PAM+MI group and the association between loneliness and PA (227). However, as this association was not present in **study IV**, the finding could rather be accounted for either by the content of MI (supportive and emphatic listening and partnership with the counsellor) (147) or as a small difference occurring by chance in a secondary outcome. Either way, future research should seek to replicate both the findings on PA and loneliness to inform decision-makers considering how to implement PAM-based interventions in either preventive or rehabilitation programs among older adults.

Methodological considerations

This section presents a discussion of the methods of the four studies within the best practice guidelines and how their strength and limitations could potentially have affected the results.

Technological literacy and readiness

Assuming that the number of days with valid data (actual wear time of the PAM and correct synchronization) are proxy-measures of the technological readiness of the participants in **study IV**, the older participants were found to have lower technological readiness compared to the younger participants. In **study IV**, as previously mentioned, the participants who had their average daily step count imputed because they had less than seven valid days of data in the intervention period were older than the participants who had more than seven days of valid data. Similarly, the number of missing days in the intervention period increased with participant age. However, as illustrated in

Figure 17, both variables violate the assumption of normally distributed variables. This makes the non-parametric alternative a better fit for the data. As it happened, it did not show any significant correlation between the number of missing days and the participant age. Overall, the results do not allow simple conclusions to be drawn on the relationship between age and quality of the data, but as the linear correlation is positive, and the dichotomous analysis (less or more than seven days of data) did reveal an age difference, the results still highlight a need for future PAM-based studies to consider the level of technological readiness among the oldest participants. This problem has been raised, discussed, and investigated several times (230–237). Technology literacy and readiness among older adults are complex issues and while not being the main focus of the four included studies in the thesis, they are highly relevant for the applicability of the results. The relevant messages include, but are not limited to, that the use of consumer-available PAMs among older adults is feasible (232,235–237) but that initial positive attitudes of using the PAM does not guarantee adherence (232), that the need for age-relevant trackers designed for age-related impairments (vision, hearing, memory, and dexterity) is high (230) and that usability might be improved by targeting the older adults directly with simple paper-based manuals and easy synchronization with less-expensive smartphones (236). To account for these issues, **study III** and **study IV** (the only studies where participants had to use the PAMs themselves) included simple and thorough paper-based manuals made by the investigators to help with the set-up of the PAMs and telephone support to both groups, no matter the allocation. This support was also implemented by directly contacting the participants if no synchronization had been performed for seven consecutive days. However, the generalisability of the results might be affected by the selection of participants, as all participants were voluntary and all were able to use smartphones and e-mail systems, as per to the inclusion criteria. Furthermore, more than 84% of the participants included in **study IV** reported that they were motivated for increase their PA level, and 33% reported having used a PAM before. Hence, the sample of older adults might be too well-functioning in general to be representative for all older adults in need of PA-related behavioural change. Even though the results of **study IV** only hint at an association between missing days and increasing age, due to lower technological literacy and readiness, the issues should be thoroughly considered by researchers before conducting PAM-based research in older adults. Furthermore, as the results suggest, there were some issues around technological readiness and literacy in **study III** and **study IV**. Although the studies were not designed to investigate this topic specifically, it should still be considered. There is no direct evidence of attrition bias (difference in loss of data or participants between the groups), which means that it applies for both groups and that the applicability of results from **study IV** might also be affected.

As these issues are probably targeted best by qualitative or mixed-method investigations, the participant-perspective and feasibility of the experimental intervention should be included in the design of the studies, which was beyond the scope of present PhD project because of time constraints.

Intervention reporting and replication

The knowledge and results from the four studies can and should be used if implementing PAM based interventions to enhance PA levels among older adults. Even though the results of **study IV** are limited by statistical imprecision, the recommendations for clinical practice, described at the end of the thesis, are still valid. However, for future research, replication of study results and methods are important. Following the PREPARE checklist on conducting a RCT, the Better Reporting of Interventions: Template for Intervention Description and Replication (TIDieR) checklist and guide is recommended to describe the intervention (161,238). While the TIDieR checklist was used to describe the intervention content in **study IV**, an MI intervention is nevertheless personally tailored and highly dependent on the relationship between the individual participant and the MI counsellor (145). This increases the difficulty of direct replication, which is a limitation to the applicability of **study IV**. Nonetheless, adherence and fidelity of the MI intervention was assessed, and future investigations should use the results from the Motivational Interviewing Treatment Integrity Scale domains (cultivating change talk, softening sustain talk, partnership, and empathy) as well as the reflection to question ratio to compare the intervention fidelity between studies (186). The Trans Theoretical Model and the Stages of Change were assessed by the MI counsellor at every phone call. However, as these assessments were not performed systematically and only used to guide the MI counsellor, they should not be used to make comparisons between participants or with participants from other studies. In summary, the level of motivation and readiness for behaviour change among individual participants should not be compared between studies. Rather, what can be compared is the fidelity of the intervention, though only to the extent that the duration and frequency of the MI calls can be replicated.

Multiple imputations to account for missing data

Because a PAM will only measure PA when being worn, no estimation of PA levels among a study sample will be complete and without missing data. Overall in **study IV**, the median days of missing data because of non-wear (<100 daily steps measured) was 5 days out of 84 possible; however, 25% of the participants had missing data for 29 or more days, which highlights presence of missing data in **study IV**, even though the amount of missing data did not differ between the groups. Compared to data cleaning, harmonisation, and imputation in accelerometery (239), data imputation in consumer-available PAMs is more simple. As the data were limited to daily step counts and the step detection algorithms in the PAMs were not accessible, the imputations of **study IV** focused on representative daily step counts and not on specific time slots of the day. Consequently, it was chosen to impute missing data from within-participant data, based on the principle that the individual average step count (if more than seven valid days were obtained during the intervention period) was the best method. The major limitation to this is situations where the participants experienced a period of lower levels of PA due to adverse events, illness or personal factors (as listed in the Harms section of **study IV**) and chose not to wear the PAM, despite having been instructed to wear it for all waking hours. This might be categorised and explained by social desirability bias and possible non-response bias. However, as seen in the results section of **study IV**, there was no difference in the number of missing days between the PAM+MI group and the PAM group. Furthermore, as seen in the post-hoc calculations section, there were no differences between participants who reported an adverse event, illness, or similar and the participants who did not. This means that the above-mentioned situation remains hypothetical and unlikely to have introduced a systematic bias in the multiple imputations and hence the main findings of **study IV**.

Intention-to-treat analyses have been reported to often be inadequately described and inadequately applied in RCTs (240). This principle was applied in **study IV,** which ensured that all randomised participants were included in the analysis by performing multiple imputations with predictive mean matching. The variables included in the multiple imputations were pre-specified and included step count in the baseline week, age, and sex. To avoid cherry-picking data, the first multiple imputations were used (out of the five runs). However, a limitation to this is the assumption that the missingness was completely random (241), which was later questioned as the participants included in the multiple imputation were older compared to the participants who were not included, and hence the missing data may not be missing at random (242). Consequently, and because the missing data only accounted for approximately 5% (4/70) of the total sample, a case could have been made to use another approach to handle the data. After investigating these potential approaches, the complete-case analysis and the intention-to-treat with last observation carried forward were found to be similar to the results obtained from the multiple imputations. Thus, using a complete-case or a last observation carried forward intention-to-treat analysis would not have changed the conclusions of **study IV,** even if it would seem to have been an easier and more straightforward strategy in retrospect. Ultimately, the multiple imputation strategy was decided upon before enrolling any participants and should hence be used as the primary method.

Accordingly, the main result of **study IV** was dependent on the multiple imputations. In simple terms, the approach imputes random values based on a Gaussian distribution, which means that the results will vary each time an imputation is run. It is unfortunately common to focus on whether a result is significant and after that, how large it is (243–245). The main finding of **study IV** was a between-group difference of 909 steps per day, p=0.07. This borderline significant result is highly dependent on multiple imputations. To illustrate the problem, Figure 15 illustrates the density of p-values derived from 5000 runs of the multiple imputations, where 34% of the obtained values fell below the alpha level of 0.05 and 62% fell below the first obtained p-value of 0.07. Thus, if the conclusion on **study IV** depended on the p-value alone, the conclusion would have varied a third of the times simply because of randomness. This highlights the problem with p-values, and following the recommendations of American Statistical Association, the conclusions should not be based on p-values and thresholds (245). The effect size should be interpreted with an indication of its estimation accuracy – in this case, 95% confidence intervals – to be transparently available for evaluation for its importance (244). From inspecting the confidence interval from the effect size, it is clear that with this imputation run, the effect most likely lies somewhere between 71 steps in favour of the PAM group and 1,889 steps in favour of the PAM+MI group. Consequently, the effect might be clinically relevant, but the imprecision of the results limits the applicability and hence no simple recommendation on the use of MI among older adults can be made without investigating this further and obtaining a more precise estimate.

Biases affecting the studies and physical activity monitoring in general

The Hawthorne effect, social desirability, blinding and performance bias

The Hawthorne effect refers to a proposed apprehension bias where people are expected to behave differently as a result of their awareness of being monitored (246). The original Hawthorne studies, initiated in 1924, reported changes in production among factory workers in Illinois in a study of how the lighting affected work practice that ultimately happened merely as a result of participation in the study, not because of the intervention itself (247). The Hawthorne effect has been proposed as an explanation of intervention effects in many research areas, and the general conception is that the study participants, reacting to their beliefs about the researchers' expectations, will modify their behaviour similarly to how social desirability and conformity bias (where participants behave as how they expect the other participants to behave, instead of behaving as they would normally do) can affect study results (246). In the thesis, **study IV** could have been affected by the Hawthorne effect. However, the Hawthorne effect would have affected the behaviour of the participants in both groups, as both groups were being monitored by PAMs. This limits the possible bias in the results, as the primary analyses focused on the difference between groups that both received an intervention where participants in both groups knew they were being monitored. Furthermore, a systematic review by McCambridge et al. reports that the Hawthorne effect is a highly complex construct to evaluate and calls for new specific concepts, as the mechanisms and magnitudes of the Hawthorne effect in health sciences are not uniform (246). There is no simple explanation for the phenomenon, nor should we expect to observe a simple impact on the results in any one direction.

As also discussed in Paper F, the full study report of **study IV**, PA research should use measures consequently throughout the period of interest in order to measure the average intervention behaviour of the participants, not only at the beginning and end of a study. The vast majority of experimental MI PA behaviour research in older adults uses a baseline and end-point week design (64–66,248,249). This type of design is more prone to bias such as those mentioned above, but also performance bias, as the nature of the interventions does not allow for blinding of the participants. However, performance bias due to knowledge of the allocation might also have affected the results of **study IV** and caused an overestimation of the effect. Although both groups received an intervention, the participants in the PAM group knew that they did not receive MI, as the study aim had been thoroughly explained in the informed consent material, which is a limitation of the study design. A systematic review by Savovic et al. investigated the effect of double blinding (both researchers and participants) in 1,973 trials and reported an overestimation of effect size of approximately 13% in studies without double-blinding and that studies with subjective outcome assessments were more prone to bias (250). However, the design of most behavioural change interventions does not allow for participants to be blinded. In many cases, it is also not necessary to obtain a scientifically sound result, either because the participants are not expected to behave differently if the allocation concealment is broken, or if the financial cost of blinding would compromise other methodological aspects of the intervention (251). In sum, it was not possible to blind the participants of **study IV** because enforced blinding would have compromised the relationship between the participant and the MI counsellor. On the other hand, detection bias from the outcome assessor's knowledge about allocation is not present in the primary outcome of **study IV,** as daily steps were objectively measured, handled automatically by the software and analysed by a blinded investigator, which is similar to most of the studies included in **study I**, and one of the strengths of using PAMs in behavioural change research.

In **Study III**, participants also knew they were being monitored, and thus could have changed their behaviour. However, social desirability and recall bias could similarly have affected the response and completion of the PAQs, which together could explain some of the inadequate validity that was ultimately found. Studies have nevertheless reported social desirability not to affect self-reported PA *per se* (252), while recall bias is still suggested to be an important main barrier to the validity of self-reported PAQs (87,253). The data of **Study III** do not allow for investigation of how specific biases could affect the results. Either way, no matter what type of bias affected the results, the validity of the PAQs was inadequate, and the results and conclusion still apply, probably to some extent because of some of the mentioned biases.

Minimal clinically important difference and statistical power

When the power calculation for **study IV** was performed, no minimal clinically important difference was applied, as there was insufficient and no conclusive evidence to specify this. Consequently, the generic rule-of-thumb interpretation for effect sizes was used, with the limitations that apply for a generic statement (24). A priori, **study IV** was planned to be powered to investigate a moderate effect between the groups, equivalent to half a standard deviation. As previously mentioned, differences in 1,000 steps per day, especially in sedentary populations or in populations who do not meeting the global recommendations for PA, seem both clinically important (17,22,140) and especially relevant to this population of interest. Future MI studies should be powered to investigate that effect size, as the effect of MI among older adults is still unclear. With the median standard deviation extracted from the results of the studies included in **study I** being 2,620 steps, future MI-randomised controlled trials should aim to include 218 participants in total (to show a difference of 1,000 steps between two groups of participants with an expected standard deviation of 2,620 steps, an alpha level of 0.05, and a power-threshold of 80%). In comparison, only four studies of the 21 included in **study I** met that threshold (45,203,207,209). Even though the process of conducting such large-scale MI trials might be arduous, as this PhD project clearly shows, aiming for 218 allocated participants should be considered best-practice for future projects to avoid wasting research on underpowered trials.

Considering the above, the results from **study IV** should rather guide future research than inform clinical decision-makers about whether or not to include an MI-intervention among older adults using PAMs for increasing their daily PA levels.

CONCLUSIONS

As previously mentioned, population ageing and inadequate PA levels among many older adults highlight the need for interventions to promote PA to be upscaled. This PhD project included four main studies to answer how PA monitoring could be used to enhance PA levels among older adults. The systematic review and meta-analysis in **study I** found low quality of evidence for a clinically highly relevant, moderate effect from PAM-based interventions on PA among older adults, which highlights the potential in using PAMs as behavioural change facilitators. The validation studies, **study II** and **study III**, found hip-worn consumer-available PAMs to be valid as outcome instruments in a population of older adults with and without assistive devices and found PAQs to perform inadequately, as they did not reflect the objectively measured PA. Finally, **study IV** found MI to be potentially effective and feasible when added to a PAM-based intervention among Danish community-dwelling older adults. However, as **study IV** was underpowered, further research should be conducted to replicate the findings with adequate sample sizes. In summary, and to answer how PA monitoring can be used to enhance PA levels among older adults, PAMs seem to be effectively increasing the PA level among older adults with clinically relevant results. Consumer-available monitors can be used as outcome-measurements in clinical settings, but hip-worn monitors should be prioritised if the population of older adults includes users of assistive devices. PAQs should be used with caution among older adults and, to avoid problems with the validity of the results, only if the respondents are given adequate instructions about answering the surveys. The use of behavioural change interventions such as MI in combination with PAMs seems promising but should be investigated further to inform decision-makers on how to appropriately implement PAMs among older adults.

IMPLICATIONS FOR FUTURE RESEARCH

This project raised seven specific implications for future research:

1. Future research should build upon the expected moderate effect size from PAMs on the PA level among older adults and avoid using passive comparators, such as waitlist control groups. Instead, future studies should either 1) compare PAM-based interventions with other behavioural change interventions directly, 2) use PAMs as an add-on intervention to investigate

if PAMs might increase the effect, or 3) use other behavioural change interventions to enhance the effect of PAMs among older adults.

2. To avoid waste of research funds by conducting underpowered trials, future two-arm RCTs should include more than 218 participants to investigate differences in 1,000 daily steps with expected standard deviations of 2,620 daily steps. This applies to all studies investigating behavioural change related to PA measured by daily step counts.
3. Consumer-available wrist-worn PAMs should only be used among older adults with normal gait patterns. If future studies include a broad and more general population of older adults, where slow gaits and use of assistive devices are to be expected, researchers should use hip-worn PAMs to increase validity.
4. Future studies should avoid using general PAQs to measure PA levels among older adults without a thorough investigation of the feasibility and validity of the instruments in the population. Instead, future studies should rely on objectively measured PA or use PAQs that have been specifically adapted and validated among older adults.
5. When using consumer-available PAMs among older adults, future studies should be mindful of the technological literacy among the study sample. Especially among the oldest participants, researchers should be ready to provide further guidance on the wear and synchronization of the PAMs.
6. When investigating the effect of assignment to the intervention, researchers should use the most appropriate imputation method to account for the inevitable missing data, whether that means using within-participant data or applying multiple imputation methods. If multiple imputation methods are chosen, researchers should perform appropriate sensitivity analyses and be aware that the obtained result is dependent on random factors and will vary between every imputation run.
7. MI serves as a relevant and feasible add-on intervention to PAM-based interventions. Future well-powered studies should replicate the design of the MIPAM intervention to obtain a more precise estimate of the effect size.

IMPLICATIONS FOR PRACTICE

This project has several implications for future health or social practice on how to use PA monitoring among older adults. Overall, PAMs are able to serve as effective, feasible facilitators of PA behaviour change among older adults. If decision makers, health care professionals or social workers choose to use PAMs actively in preventive programs or in rehabilitation settings, the following should be considered:

1. PAMs have shown to be effective in general. However, individual intervention participants might react differently to the feedback. Even though most users (on average) should react positively and use the feedback in goal setting to change their behaviour, some users might feel anxiety about receiving real time feedback from the PAMs, while others might have problems with usability due to low technological literacy. With this in mind, and as the latter might be associated with increasing age, PAMs should not be implemented in broader populations without proper instructions and follow up from health or social workers.
2. Consumer-available PAMs can be used in similar ways as research-grade PAMs, as they both have been found to be reliable in measuring and quantifying steps among older adults. Furthermore, consumer-available PAMs might be favourable to some, as the devices are constantly renewed and updated to improve consumer-experience. However, if consumer-available PAMs are used among users who are expected to use assistive devices, hip-worn PAMs are to be preferred.
3. General PAQs should not be trusted as valid measurement instruments among older adults due to problems with the interpretation and completion of the surveys.
4. Not all consumer-available PAMs are relevant for older adults. To avoid usability problems, PAMs with a long battery life, easy synchronization and simple feedback (e.g. on steps) should be chosen over advanced fitness or sports trackers.
5. If steps are chosen as the PA construct of interest, users should be educated in the association between daily steps and healthy aging. Practitioners should explain the emerging evidence suggesting highly relevant health benefits of increasing activity levels even with 1,000 daily steps, as unobtainable step counts (e.g. 10,000/day) might affect the users negatively.
6. MI seems to be an effective and feasible way of increasing the effect from PAMs. However, as further research with adequately powered randomised controlled trials is needed to establish precise effect size estimates, MI might still be unnecessary and should not be implemented in combination with PAMs at this point to avoid gratuitous spending of limited funds.

REFERENCES

1. World Health Organization. Global strategy and action plan on ageing and health [Internet]. 2017. Available from: http://www.who.int/ageing/WHO-GSAP-2017.pdf?ua=1

2. WHO. WHO | What is Healthy Ageing? [Internet]. What is Healthy Ageing? World Health Organization; 2020 [cited 2020 Jun 15]. Available from: http://www.who.int/ageing/healthy-ageing/en/

3. Daskalopoulou C, Stubbs B, Kralj C, Koukounari A, Prince M, Prina AM. Physical activity and healthy ageing: A systematic review and meta-analysis of longitudinal cohort studies. Ageing Res Rev [Internet]. 2017 Sep 1 [cited 2020 Jun 15];38:6–17. Available from: http://www.sciencedirect.com/science/article/pii/S1568163717300302

4. Bosnes I, Nordahl HM, Stordal E, Bosnes O, Myklebust TÅ, Almkvist O. Lifestyle predictors of successful aging: A 20-year prospective HUNT study. PLOS ONE [Internet]. 2019 Nov 7 [cited 2020 Jun 15];14(7):e0219200. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0219200

5. European Commission. The 2018 Ageing Report Economic & Budgetary Projections for the 28 EU Member States (2016-2070) [Internet]. European Economy Institutional Papers; 2018. Available from: https://ec.europa.eu/info/sites/info/files/economy-finance/ip079_en.pdf

6. World Bank. Life expectancy at birth, total (years) | Data [Internet]. Life expectancy at birth, total (years). 2020 [cited 2020 Jun 15]. Available from: https://data.worldbank.org/indicator/SP.DYN.LE00.IN?end=2018&start=2000

7. World Bank. Population ages 65 and above (% of total population) - Denmark | Data [Internet]. Population ages 65 and above (% of total population). 2020 [cited 2020 Jun 15]. Available from: https://data.worldbank.org/indicator/SP.POP.65UP.TO.ZS?locations=DK&view=chart

8. Forouzanfar MH, Afshin A, Alexander LT, Anderson HR, Bhutta ZA, Biryukov S, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. The Lancet [Internet]. 2016 Oct 8 [cited 2020 Jun 15];388(10053):1659–724. Available from: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(16)31679-8/abstract

9. WHO. Non communicable diseases [Internet]. Noncommunicable diseases. 2020 [cited 2020 Jun 15]. Available from: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases

10. Lee I-M, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Impact of Physical Inactivity on the World's Major Non-Communicable Diseases. Lancet [Internet]. 2012 Jul 21 [cited 2020 Jun 15];380(9838):219–29. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3645500/

11. Bangsbo J, Blackwell J, Boraxbekk C-J, Caserotti P, Dela F, Evans AB, et al. Copenhagen Consensus statement 2019: physical activity and ageing. Br J Sports Med [Internet]. 2019 Jul 1 [cited 2020 Feb 11];53(14):856–8. Available from: https://bjsm.bmj.com/content/53/14/856

12. Gomes M, Figueiredo D, Teixeira L, Poveda V, Paúl C, Santos-Silva A, et al. Physical inactivity among older adults across Europe based on the SHARE database. Age Ageing [Internet]. 2017 Jan 19 [cited 2018 Feb 6];46(1):71–7. Available from: https://academic.oup.com/ageing/article/46/1/71/2281655

13. WHO | The Global Strategy on Diet, Physical Activity and Health (DPAS) [Internet]. WHO. [cited 2019 Jan 17]. Available from: http://www.who.int/nmh/wha/59/dpas/en/

14. Danish health Authority. Danskernes Sundhed – Den Nationale Sundhedsprofil 2017. Sundhedsstyrelsen; 2018.

15. Ding D, Lawson KD, Kolbe-Alexander TL, Finkelstein EA, Katzmarzyk PT, van Mechelen W, et al. The economic burden of physical inactivity: a global analysis of major non-communicable diseases. The Lancet [Internet]. 2016 Sep [cited 2020 Jun 15];388(10051):1311–24. Available from: https://linkinghub.elsevier.com/retrieve/pii/S014067361630383X

16. Torjesen I. Global cost of physical inactivity is estimated at $67.5bn a year. BMJ [Internet]. 2016 Jul 28 [cited 2020 Jul 3];354. Available from: https://www.bmj.com/content/354/bmj.i4187

17. Hansen BH, Dalene KE, Ekelund U, Fagerland MW, Kolle E, Steene-Johannessen J, et al. Step by Step: Association of Device-Measured Daily Steps With All Cause-Mortality – a Prospective Cohort Study. Scand J Med Sci Sports [Internet]. 2020 [cited 2020 May 27];n/a(n/a). Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/sms.13726

18. Ekelund U, Tarp J, Steene-Johannessen J, Hansen BH, Jefferis B, Fagerland MW, et al. Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. BMJ [Internet]. 2019 Aug 21 [cited 2020 Jul 21];366. Available from: https://www.bmj.com/content/366/bmj.l4570

19. Taylor D. Physical activity is medicine for older adults. Postgrad Med J [Internet]. 2014 Jan 1 [cited 2020 Jun 25];90(1059):26–32. Available from: https://pmj.bmj.com/content/90/1059/26

20. Taylor AH, Cable NT, Faulkner G, Hillsdon M, Narici M, Van Der Bij AK. Physical activity and older adults: a review of health benefits and the effectiveness of interventions. J Sports Sci. 2004 Aug;22(8):703–25.

21. Mok A, Khaw K-T, Luben R, Wareham N, Brage S. Physical activity trajectories and mortality: population based cohort study. BMJ [Internet]. 2019 Jun 26 [cited 2019 Jul 31];l2323. Available from: http://www.bmj.com/lookup/doi/10.1136/bmj.l2323

22. Lee I-M, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of Step Volume and Intensity With All-Cause Mortality in Older Women. JAMA Intern Med [Internet]. 2019 Aug 1 [cited 2020 Mar 20];179(8):1105–12. Available from: https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2734709

23. Szanton SL, Walker RK, Roberts L, Thorpe RJ, Wolff J, Agree E, et al. Older adults' favorite activities are resoundingly active: Findings from the NHATS study. Geriatr Nur (Lond) [Internet]. 2015 Mar [cited 2018 Mar 16];36(2):131–5. Available from: http://linkinghub.elsevier.com/retrieve/pii/S0197457214004133

24. Higgins JPT, Green S. 12.6.2 Re-expressing SMDs using rules of thumb for effect sizes. In: Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 2011 [cited 2020 May 20]. Available from: https://handbook-5-1.cochrane.org/chapter_12/12_6_2_re_expressing_smds_using_rules_of_thumb_for_effect_sizes.htm

25. Grande GD, Oliveira CB, Morelhão PK, Sherrington C, Tiedemann A, Pinto RZ, et al. Interventions Promoting Physical Activity Among Older Adults: A Systematic Review and Meta-Analysis. The Gerontologist [Internet]. 2019 [cited 2020 Jun 16]; Available from: https://academic.oup.com/gerontologist/advance-article/doi/10.1093/geront/gnz167/5685409

26. Reis RS, Salvo D, Ogilvie D, Lambert EV, Goenka S, Brownson RC, et al. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. Lancet Lond Engl. 2016 Sep 24;388(10051):1337–48.

27. World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world. Geneva: World Health Organization; 2018.

28. Elsawy B, Higgins KE. Physical activity guidelines for older adults. Am Fam Physician. 2010 Jan 1;81(1):55–9.

29. WHO. WHO | Physical Activity and Older Adults [Internet]. WHO. 2011 [cited 2017 Dec 18]. Available from: http://www.who.int/dietphysicalactivity/factsheet_olderadults/en/

30. Murtagh EM, Murphy MH, Boone-Heinonen J. Walking: the first steps in cardiovascular disease prevention. Curr Opin Cardiol. 2010 Sep;25(5):490–6.

31. Chodzko-zajko WJ, Proctor DN, Singh MAF, Minson CT, Nigg CR, Salem GJ, et al. Exercise and Physical Activity for Older Adults. Med Sci Sports Exerc [Internet]. 2009 Jul 1 [cited 2017 Dec 18];41(7):1510–30. Available from: https://insights.ovid.com/pubmed?pmid=19516148

32. Dimitrov DV. Medical Internet of Things and Big Data in Healthcare. Healthc Inform Res [Internet]. 2016 Jul [cited 2018 Jun 7];22(3):156–63. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4981575/

33. Piwek L, Ellis DA, Andrews S, Joinson A. The Rise of Consumer Health Wearables: Promises and Barriers. PLOS Med [Internet]. 2016 Feb 2 [cited 2017 Oct 23];13(2):e1001953. Available from: http://dx.plos.org/10.1371/journal.pmed.1001953

34. Stunkard A. A Method of Studying Physical Activity in Man. Am J Clin Nutr [Internet]. 1960 Sep 1 [cited 2017 Dec 20];8(5):595–601. Available from: http://ajcn.nutrition.org/content/8/5/595

35. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA. 2015 Feb 3;313(5):459–60.

36. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA. 2007 Nov 21;298(19):2296–304.

37. Koizumi D, Rogers NL, Rogers ME, Islam MM, Kusunoki M, Takeshima N. Efficacy of an accelerometer-guided physical activity intervention in community-dwelling older women. J Phys Act Health. 2009 Jul;6(4):467–74.

38. McLellan AG, Slaght J, Craig CM, Mayo A, Sénéchal M, Bouchard DR. Can older adults improve the identification of moderate intensity using walking cadence? Aging Clin Exp Res. 2018 Jan;30(1):89–92.

39. Mutrie N., Doolin O., Fitzsimons CF., Grant PM., Granat M., Grealy M., et al. Increasing older adults' walking through primary care: results of a pilot randomized controlled trial. Fam Pract. 2012;29(6):633–42.

40. Nishiguchi S, Yamada M, Tanigawa T, Sekiyama K, Kawagoe T, Suzuki M, et al. 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. 2015 Jul;63(7):1355–63.

41. Rowley TW., Lenz EK., Swartz AM., Miller NE., Maeda H., Strath SJ. Efficacy of an Individually Tailored, Internet-Mediated Physical Activity Intervention in Older Adults: A Randomized Controlled Trial. J Appl Gerontol Off J South Gerontol Soc. 2017;733464817735396.

42. Tabak M., Vollenbroek-Hutten MM., van der Valk PD., van der Palen J., Hermens HJ. A telerehabilitation intervention for patients with Chronic Obstructive Pulmonary Disease: a randomized controlled pilot trial. Clin Rehabil. 2014;28(6):582–91.

43. Croteau KA, Richeson NE, Vines SW, Jones DB. Effects of a pedometer-based physical activity program on older adults' mobility-related self-efficacy and physical performance. Act Adapt Aging [Internet]. 2004;28(2):19–33. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=106659778&site=ehost-live

44. Kawagoshi A, Kiyokawa N, Sugawara K, Takahashi H, Sakata S, Satake M, et al. Effects of low-intensity exercise and home-based pulmonary rehabilitation with pedometer feedback on physical activity in elderly patients with chronic obstructive pulmonary disease. Respir Med. 2015 Mar;109(3):364–71.

45. Kolt GS, Schofield GM, Kerse N, Garrett N, Ashton T, Patel A. Healthy steps trial: pedometer-based advice and physical activity for low-active older adults. Ann Fam Med [Internet]. 2012;10(3):206–12. Available from: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=104446560&site=ehost-live

46. McMurdo MET, Sugden J, Argo I, Boyle P, Johnston DW, Sniehotta FF, et al. Do pedometers increase physical activity in sedentary older women? A randomized controlled trial. J Am Geriatr Soc. 2010 Nov;58(11):2099–106.

47. Sugden J.A., Sniehotta F.F., Donnan P.T., Boyle P., Johnston D.W., McMurdo M.E.T. The feasibility of using pedometers and brief advice to increase activity in sedentary older women - A pilot study. BMC Health Serv Res. 2008;8((Boyle) School of Geography and Geosciences, University of St Andrews, Irvine Building, North Street, St Andrews, United Kingdom):169.

48. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med Publ Soc Behav Med. 2013 Aug;46(1):81–95.

49. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol Off J Div Health Psychol Am Psychol Assoc. 2009 Nov;28(6):690–701.

50. Victor CR, Rogers A, Woodcock A, Beighton C, Cook DG, Kerry SM, et al. What factors support older people to increase their physical activity levels? An exploratory analysis of the experiences of PACE-Lift trial participants. Arch Gerontol Geriatr [Internet]. 2016 [cited 2019 Feb 27];67:1–6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5019110/

51. Bull ER, McCleary N, Li X, Dombrowski SU, Dusseldorp E, Johnston M. Interventions to Promote Healthy Eating, Physical Activity and Smoking in Low-Income Groups: a Systematic Review with Meta-Analysis of Behavior Change Techniques and Delivery/Context. Int J Behav Med. 2018 Dec;25(6):605–16.

52. Dusseldorp E, van Genugten L, van Buuren S, Verheijden MW, van Empelen P. Combinations of techniques that effectively change health behavior: evidence from Meta-CART analysis. Health Psychol Off J Div Health Psychol Am Psychol Assoc. 2014 Dec;33(12):1530–40.

53. Knittle K, Nurmi J, Crutzen R, Hankonen N, Beattie M, Dombrowski SU. How can interventions increase motivation for physical activity? A systematic review and meta-analysis. Health Psychol Rev. 2018;12(3):211–30.

54. Miller WR, Rollnick S. Motivational interviewing: Helping people change, 3rd edition. New York, NY, US: Guilford Press; 2013. xii, 482 p. (Motivational interviewing: Helping people change, 3rd edition).

55. Cummings SM, Cooper RL, Cassie KM. Motivational Interviewing to Affect Behavioral Change in Older Adults. Res Soc Work Pract [Internet]. 2009 Mar 1 [cited 2018 Dec 17];19(2):195–204. Available from: https://doi.org/10.1177/1049731508320216

56. Hardcastle SJ, Hancox J, Hattar A, Maxwell-Smith C, Thøgersen-Ntoumani C, Hagger MS. Motivating the unmotivated: how can health behavior be changed in those unwilling to change? Front Psychol [Internet]. 2015 Jun 16 [cited 2019 Mar 7];6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4468355/

57. Municipality of Odense. Odense Kommune sender 3500 medarbejdere på uddannelse i... [Internet]. The Municipality of Odense educates 3500 employees in Motivational Interviewing. 2015 [cited 2020 Jun 23]. Available from: https://www.odense.dk/presse/pressemeddelelser/pressemeddelelser-2014/odense-kommune-sender-3500-medarbejdere-paa-uddannelse-i-motiverende-bogersamtale

58. Rodnick JF. Denmark: teaching GPs to use motivational interviewing. Fam Med. 2007 Mar;39(3):215.

59. Rubak S, Sandbæk A, Lauritzen T, Christensen B. Motivational interviewing: a systematic review and meta-analysis. Br J Gen Pract [Internet]. 2005 Apr 1 [cited 2018 Dec 16];55(513):305. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1463134/

60. Lindhardt CL, Rubak S, Mogensen O, Hansen HP, Goldstein H, Lamont RF, et al. Healthcare professionals experience with motivational interviewing in their encounter with obese pregnant women. Midwifery. 2015 Jul;31(7):678–84.

61. Charles M, Bruun NH, Simmons R, Dalsgaard E-M, Witte D, Jorgensen M, et al. The effect of training GPs in motivational interviewing on incident cardiovascular disease and mortality in people with screen-detected diabetes. Results from the ADDITION-Denmark randomised trial. BJGP Open. 2020;4(1).

62. Rosenbek Minet LK, Wagner L, Lønvig EM, Hjelmborg J, Henriksen JE. The effect of motivational interviewing on glycaemic control and perceived competence of diabetes self-management in patients with type 1 and type 2 diabetes mellitus after attending a group education programme: a randomised controlled trial. Diabetologia. 2011 Jul;54(7):1620–9.

63. Brandt CJ, Søgaard GI, Clemensen J, Sndergaard J, Nielsen JB. General Practitioners' Perspective on eHealth and Lifestyle Change: Qualitative Interview Study. JMIR MHealth UHealth. 2018 Apr 17;6(4):e88.

64. Brodie DA, Inoue A. Motivational interviewing to promote physical activity for people with chronic heart failure. J Adv Nurs [Internet]. 2005 [cited 2020 May 7];50(5):518–27. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2648.2005.03422.x

65. O'Halloran PD, Shields N, Blackstock F, Wintle E, Taylor NF. Motivational interviewing increases physical activity and self-efficacy in people living in the community after hip fracture: a randomized controlled trial. Clin Rehabil [Internet]. 2016 Nov [cited 2020 May 7];30(11):1108–19. Available from: http://journals.sagepub.com/doi/10.1177/0269215515617814

66. Audsley S, Kendrick D, Logan P, Jones M, Orton E. A randomised feasibility study assessing an intervention to keep adults physically active after falls management exercise programmes end. Pilot Feasibility Stud [Internet]. 2020 Dec [cited 2020 May 7];6(1):37. Available from: https://pilotfeasibilitystudies.biomedcentral.com/articles/10.1186/s40814-020-00570-9

67. Wright SP, Hall Brown TS, Collier SR, Sandberg K. How consumer physical activity monitors could transform human physiology research. Am J Physiol - Regul Integr Comp Physiol [Internet]. 2017 Mar 1 [cited 2018 Jun 7];312(3):R358–67. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5401997/

68. Cadmus-Bertram L. Using Fitness Trackers in Clinical Research: What Nurse Practitioners Need to Know. J Nurse Pract JNP. 2017 Jan;13(1):34–40.

69. Dowd KP, Szeklicki R, Minetto MA, Murphy MH, Polito A, Ghigo E, et al. A systematic literature review of reviews on techniques for physical activity measurement in adults: a DEDIPAC study. Int J Behav Nutr Phys Act [Internet]. 2018 Feb 8 [cited 2020 Mar 22];15(1):15. Available from: https://doi.org/10.1186/s12966-017-0636-2

70. Lee I, Park S. A Comparison of Gait Characteristics in the Elderly People, People with Knee Pain, and People Who Are Walker Dependent People. J Phys Ther Sci [Internet]. 2013 Aug [cited 2018 Apr 4];25(8):973–6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3820225/

71. Floegel TA, Florez-Pregonero A, Hekler EB, Buman MP. Validation of Consumer-Based Hip and Wrist Activity Monitors in Older Adults With Varied Ambulatory Abilities. J Gerontol A Biol Sci Med Sci. 2017 Feb;72(2):229–36.

72. Farina N, Lowry RG. The Validity of Consumer-Level Activity Monitors in Healthy Older Adults in Free-Living Conditions. J Aging Phys Act [Internet]. 2018 Jan 1;26(1):128–35. Available from: https://journals.humankinetics.com/doi/pdf/10.1123/japa.2016-0344

73. Paul SS, Tiedemann A, Hassett LM, Ramsay E, Kirkham C, Chagpar S, et al. Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults. BMJ Open Sport Exerc Med [Internet]. 2015 Jul 1 [cited 2017 Oct 23];1(1):e000013. Available from: http://bmjopensem.bmj.com/content/1/1/e000013

74. Phillips LJ, Petroski GF, Markis NE. A Comparison of Accelerometer Accuracy in Older Adults. Res Gerontol Nurs. 2015 Oct;8(5):213–9.

75. Thorup CB, Andreasen JJ, Sørensen EE, Grønkjær M, Dinesen BI, Hansen J. Accuracy of a step counter during treadmill and daily life walking by healthy adults and patients with cardiac disease. BMJ Open [Internet]. 2017 Mar 1 [cited 2018 Nov 2];7(3):e011742. Available from: https://bmjopen.bmj.com/content/7/3/e011742

76. Boeselt T, Spielmanns M, Nell C, Storre JH, Windisch W, Magerhans L, et al. Validity and Usability of Physical Activity Monitoring in Patients with Chronic Obstructive Pulmonary Disease (COPD). PLOS ONE [Internet]. 2016 Jun 15 [cited 2018 Nov 2];11(6):e0157229. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0157229

77. Alharbi M, Bauman A, Neubeck L, Gallagher R. Validation of Fitbit-Flex as a measure of free-living physical activity in a community-based phase III cardiac rehabilitation population. Eur J Prev Cardiol [Internet]. 2016 Sep [cited 2018 Nov 2];23(14):1476–85. Available from: http://journals.sagepub.com/doi/10.1177/2047487316634883

78. Straiton N, Alharbi M, Bauman A, Neubeck L, Gullick J, Bhindi R, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: A systematic review. Maturitas [Internet]. 2018 Jun 1 [cited 2018 Nov 2];112:85–93. Available from: http://www.sciencedirect.com/science/article/pii/S0378512218301828

79. Westerterp KR. Assessment of physical activity: a critical appraisal. Eur J Appl Physiol [Internet]. 2009 Apr 1 [cited 2020 Mar 18];105(6):823–8. Available from: https://doi.org/10.1007/s00421-009-1000-2

80. Doherty A, Jackson D, Hammerla N, Plötz T, Olivier P, Granat MH, et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. PloS One. 2017;12(2):e0169649.

81. Pedišić Ž, Bauman A. Accelerometer-based measures in physical activity surveillance: current practices and issues. Br J Sports Med [Internet]. 2015 Feb 1 [cited 2020 Mar 17];49(4):219–23. Available from: https://bjsm.bmj.com/content/49/4/219

82. Ku P-W, Steptoe A, Liao Y, Hsueh M-C, Chen L-J. A Threshold of Objectively-Assessed Daily Sedentary Time for All-Cause Mortality in Older Adults: A Meta-Regression of Prospective Cohort Studies. J Clin Med [Internet]. 2019 Apr 25 [cited 2020 Jun 18];8(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6517908/

83. Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. Res Q Exerc Sport. 2000 Jun;71 Suppl 2:1–14.

84. Colley RC, Butler G, Garriguet D, Prince SA, Roberts KC. Comparison of self-reported and accelerometer-measured physical activity in Canadian adults. Health Rep. 2018 19;29(12):3–15.

85. Helmerhorst HJF, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. Int J Behav Nutr Phys Act. 2012 Aug 31;9:103.

86. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act. 2011 Oct 21;8:115.

87. Sagelv EH, Hopstock LA, Johansson J, Hansen BH, Brage S, Horsch A, et al. Criterion validity of two physical activity and one sedentary time questionnaire against accelerometry in a large cohort of adults and older adults. BMJ Open Sport — Exerc Med [Internet]. 2020 Feb 26 [cited 2020 Mar 15];6(1). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7047487/

88. Cleland C, Ferguson S, Ellis G, Hunter RF. Validity of the International Physical Activity Questionnaire (IPAQ) for assessing moderate-to-vigorous physical activity and sedentary behaviour of older adults in the United Kingdom. BMC Med Res Methodol [Internet]. 2018 Dec 22 [cited 2020 Mar 16];18. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6303992/

89. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc. 2003 Aug;35(8):1381–95.

90. Danquah IH, Petersen CB, Skov SS, Tolstrup JS. Validation of the NPAQ-short – a brief questionnaire to monitor physical activity and compliance with the WHO recommendations. BMC Public Health. 2018 May 8;18(1):601.

91. Danquah IH, Petersen CB, Skov SS, Tolstrup JS. Validation of the NPAQ-short – a brief questionnaire to monitor physical activity and compliance with the WHO recommendations. BMC Public Health [Internet]. 2018 May 8 [cited 2018 Dec 28];18(1):601. Available from: https://doi.org/10.1186/s12889-018-5538-y

92. Silsbury Z, Goldsmith R, Rushton A. Systematic review of the measurement properties of self-report physical activity questionnaires in healthy adult populations. BMJ Open. 2015 Sep 15;5(9):e008430.

93. Kolbe-Alexander TL, Lambert EV, Harkins JB, Ekelund U. Comparison of two methods of measuring physical activity in South African older adults. J Aging Phys Act. 2006 Jan;14(1):98–114.

94. Deng HB, Macfarlane DJ, Thomas GN, Lao XQ, Jiang CQ, Cheng KK, et al. Reliability and validity of the IPAQ-Chinese: the Guangzhou Biobank Cohort study. Med Sci Sports Exerc. 2008 Feb;40(2):303–7.

95. European Commission, Statistical Office of the European Union. Ageing Europe: looking at the lives of older people in the EU : 2019 edition. [Internet]. 2019 [cited 2020 Jun 30]. Available from: https://op.europa.eu/publication/manifestation_identifier/PUB_KS0219681ENN

96. United Nations. World Population Ageing 2019. Department of Economic and Social Affairs Population Division; 2020.

97. Bowling A. Aspirations for Older Age in the 21st Century: What is Successful Aging?: Int J Aging Hum Dev [Internet]. 2016 Jul 22 [cited 2020 Jun 15]; Available from: https://journals.sagepub.com/doi/10.2190/L0K1-87W4-9R01-7127

98. United Nations. United Nations: Key conference outcomes in ageing [Internet]. United Nations; 1982 [cited 2020 Jun 30]. Available from: https://www.un.org/en/development/devagenda/ageing.shtml

99. Wong RY. A New Strategic Approach to Successful Aging and Healthy Aging. Geriatrics [Internet]. 2018 Nov 29 [cited 2020 Jun 30];3(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6371117/

100. Franceschi C, Garagnani P, Morsiani C, Conte M, Santoro A, Grignolio A, et al. The Continuum of Aging and Age-Related Diseases: Common Mechanisms but Different Rates. Front Med [Internet]. 2018 Mar 12 [cited 2020 Jul 1];5. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5890129/

101. Colón-Emeric CS, Whitson HE, Pavon J, Hoenig H. Functional Decline in Older Adults. Am Fam Physician [Internet]. 2013 Sep 15 [cited 2020 Jul 1];88(6):388–94. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3955056/

102. Jin K. Modern Biological Theories of Aging. Aging Dis [Internet]. 2010 Aug 1 [cited 2020 Jul 1];1(2):72–4. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2995895/

103. Aunan JR, Watson MM, Hagland HR, Søreide K. Molecular and biological hallmarks of ageing. Br J Surg. 2016 Jan;103(2):e29-46.

104. Lipsky MS, King M. Biological theories of aging. Dis--Mon DM. 2015 Nov;61(11):460–6.

105. Wernher I, Lipsky MS. Psychological theories of aging. Dis--Mon DM. 2015 Nov;61(11):480–8.

106. Hasworth SB, Cannon ML. Social theories of aging: A review. Dis--Mon DM. 2015 Nov;61(11):475–9.

107. Rogers NT, Marshall A, Roberts CH, Demakakos P, Steptoe A, Scholes S. Physical activity and trajectories of frailty among older adults: Evidence from the English Longitudinal Study of Ageing. PLoS ONE [Internet]. 2017 Feb 2 [cited 2020 Jul 1];12(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5289530/

108. Song J, Lindquist LA, Chang RW, Semanik PA, Ehrlich-Jones LS, Lee J, et al. Sedentary Behavior as a Risk Factor for Physical Frailty Independent of Moderate Activity: Results From the Osteoarthritis Initiative. Am J Public Health. 2015 Jul;105(7):1439–45.

109. Haider S, Grabovac I, Dorner TE. Fulfillment of physical activity guidelines in the general population and frailty status in the elderly population: A correlation study of data from 11 European countries. Wien Klin Wochenschr [Internet]. 2019 Jun [cited 2020 Jul 1];131(11–12):288–93. Available from: http://link.springer.com/10.1007/s00508-018-1408-y

110. Paterson DH, Warburton DE. Physical activity and functional limitations in older adults: a systematic review related to Canada's Physical Activity Guidelines. Int J Behav Nutr Phys Act. 2010 May 11;7:38.

111. Kritikos A, Bekiari A, Nikitaras N, Famissis K, Sakellariou K. Hippocrates' counselling with regard to physical exercise, gymnastics, dietetics and health. Ir J Med Sci. 2009 Sep;178(3):377–83.

112. World Health Organization. WHO guidelines on physical activity and sedentary behaviour [Internet]. Geneva: World Health Organization; 2020. Available from: https://apps.who.int/iris/handle/10665/336656

113. Jakicic JM, Kraus WE, Powell KE, Campbell WW, Janz KF, Troiano RP, et al. Association between Bout Duration of Physical Activity and Health: Systematic Review. Med Sci Sports Exerc. 2019;51(6):1213–9.

114. World Health Organization. WHO | What is Moderate-intensity and Vigorous-intensity Physical Activity? [Internet]. WHO. 2017 [cited 2017 Jun 26]. Available from: http://www.who.int/dietphysicalactivity/physical_activity_intensity/en/

115. Barnett A, van den Hoek D, Barnett D, Cerin E. Measuring moderate-intensity walking in older adults using the ActiGraph accelerometer. BMC Geriatr [Internet]. 2016 Dec [cited 2020 Jul 15];16(1):211. Available from: http://bmcgeriatr.biomedcentral.com/articles/10.1186/s12877-016-0380-5

116. Corbett DB, Valiani V, Knaggs JD, Manini TM. Evaluating Walking Intensity with Hip-Worn Accelerometers in Elders. Med Sci Sports Exerc [Internet]. 2016 Nov [cited 2020 Jul 15];48(11):2216–21. Available from: https://journals.lww.com/00005768-201611000-00018

117. Jefferis BJ, Sartini C, Lee I-M, Choi M, Amuzu A, Gutierrez C, et al. Adherence to physical activity guidelines in older adults, using objectively measured physical activity in a population-based study. BMC Public Health [Internet]. 2014 Dec [cited 2020 Jul 14];14(1):382. Available from: http://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-14-382

118. Aguiar EJ, Ducharme SW, Thomas DM. Is 4400 Steps per Day the New 10 000 Steps per Day? JAMA Intern Med [Internet]. 2019 Nov 1 [cited 2020 Jun 29];179(11):1601–1601. Available from: https://jamanetwork.com/journals/jamainternalmedicine/fullarticle/2753704

119. Tudor-Locke C, Craig CL, Aoyagi Y, Bell RC, Croteau KA, De Bourdeaudhuij I, et al. How many steps/day are enough? For older adults and special populations. Int J Behav Nutr Phys Act [Internet]. 2011 Jul 28 [cited 2018 May 31];8:80. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3169444/

120. Sylvia LG, Bernstein EE, Hubbard JL, Keating L, Anderson EJ. A Practical Guide to Measuring Physical Activity. J Acad Nutr Diet [Internet]. 2014 Feb [cited 2020 Mar 19];114(2):199–208. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3915355/

121. N L, Gb G, R M. Measurement of Total Carbon Dioxide Production by Means of D2O18 [Internet]. Vol. 7, Journal of applied physiology. J Appl Physiol; 1955 [cited 2020 Jun 26]. Available from: https://pubmed.ncbi.nlm.nih.gov/14381351/

122. Schoeller DA, van Santen E. Measurement of energy expenditure in humans by doubly labeled water method. J Appl Physiol. 1982 Oct;53(4):955–9.

123. M.S ELMJ, Ph.D PSF, Blair DS. Physical activity assessment: A review of methods. Crit Rev Food Sci Nutr [Internet]. 1996 May 1 [cited 2020 Jun 26];36(5):385–96. Available from: https://doi.org/10.1080/10408399609527732

124. Porter J, Nguo K, Gibson S, Huggins CE, Collins J, Kellow NJ, et al. Total energy expenditure in adults aged 65 years and over measured using doubly-labelled water: international data availability and opportunities for data sharing. Nutr J [Internet]. 2018 Mar 26 [cited 2020 Mar 19];17(1):40. Available from: https://doi.org/10.1186/s12937-018-0348-8

125. Gibbs-Smith CH, Rees G. The inventions of Leonardo da Vinci. Oxford: Phaidon; 1978. 110 p.

126. Jefferson T, Wilson DL, Stanton LC. Jefferson abroad. Modern Library ed. New York: Modern Library; 1999. 338 p.

127. Wilhelm P, Perrez M, Pawlik K. Conducting research in daily life: A historical review. In: Handbook of research methods for studying daily life. New York, NY, US: The Guilford Press; 2012. p. 62–86.

128. Chirico AM, Stunkard AJ. Physical activity and human obesity. N Engl J Med. 1960 Nov 10;263:935–40.

129. Tudor-Locke C. Manpo-kei: the art and science of step counting : how to be naturally active and lose weight! Victoria, B.C.: Trafford; 2003.

130. Karas M, Bai J, Strączkiewicz M, Harezlak J, Glynn NW, Harris T, et al. Accelerometry Data in Health Research: Challenges and Opportunities. Stat Biosci [Internet]. 2019 Jul 1 [cited 2020 Mar 18];11(2):210–37. Available from: https://doi.org/10.1007/s12561-018-9227-2

131. Morris Jr. Accelerometry--a technique for the measurement of human body movements. J Biomech [Internet]. 1973 Nov 1 [cited 2020 Jun 29];6(6):729–36. Available from: https://europepmc.org/article/med/4757490

132. Cavagna G, Saibene F, Margaria R. A three-directional accelerometer for analyzing body movements. J Appl Physiol [Internet]. 1961 Jan 1 [cited 2020 Jun 29];16(1):191–191. Available from: https://journals.physiology.org/doi/abs/10.1152/jappl.1961.16.1.191

133. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act. 2015 Dec 18;12:159.

134. Gomersall SR, Ng N, Burton NW, Pavey TG, Gilson ND, Brown WJ. Estimating Physical Activity and Sedentary Behavior in a Free-Living Context: A Pragmatic Comparison of Consumer-Based Activity Trackers and ActiGraph Accelerometry. J Med Internet Res. 2016 07;18(9):e239.

135. Larsen RT, Korfitsen CB, Juhl CB, Andersen HB, Langberg H, Christensen J. Criterion validity for step counting in four consumer-grade physical activity monitors among older adults with and without rollators. Eur Rev Aging Phys Act [Internet]. 2020 Dec [cited 2020 Jan 6];17(1):1. Available from: https://eurapa.biomedcentral.com/articles/10.1186/s11556-019-0235-0

136. Sattler MC, Jaunig J, Tösch C, Watson ED, Mokkink LB, Dietz P, et al. Current Evidence of Measurement Properties of Physical Activity Questionnaires for Older Adults: An Updated Systematic Review. Sports Med Auckl Nz [Internet]. 2020 [cited 2020 Jun 30];50(7):1271–315. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7305082/

137. Eckert KG, Lange MA. Comparison of physical activity questionnaires for the elderly with the International Classification of Functioning, Disability and Health (ICF) – an analysis of content. BMC Public Health [Internet]. 2015 Mar 14 [cited 2020 Jun 30];15. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4392753/

138. van Poppel MNM, Chinapaw MJM, Mokkink LB, van Mechelen W, Terwee CB. Physical Activity Questionnaires for Adults. Sports Med [Internet]. 2010 Jul 1 [cited 2019 Mar 15];40(7):565–600. Available from: https://doi.org/10.2165/11531930-000000000-00000

139. Mai KS, Sandbaek A, Borch-Johnsen K, Lauritzen T. Are lifestyle changes achieved after participation in a screening programme for Type 2 diabetes? The ADDITION Study, Denmark. Diabet Med J Br Diabet Assoc. 2007 Oct;24(10):1121–8.

140. WHO | Global recommendations on physical activity for health [Internet]. [cited 2018 Dec 30]. Available from: https://www.who.int/dietphysicalactivity/publications/9789241599979/en/

141. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 3. ed., Pearson new internat.ed. Harlow: Pearson; 2014. 842 p. (Always Learning).

142. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010 Jul;63(7):737–45.

143. Tipton CM. The history of “Exercise Is Medicine” in ancient civilizations. Adv Physiol Educ [Internet]. 2014 Jun [cited 2020 Jul 15];38(2):109–17. Available from: https://www.physiology.org/doi/10.1152/advan.00136.2013

144. Qiu S, Cai X, Ju C, Sun Z, Yin H, Zügel M, et al. Step Counter Use and Sedentary Time in Adults. Medicine (Baltimore) [Internet]. 2015 Sep 4;94(35). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4616497/

145. Rollnick S, Miller WR. What is Motivational Interviewing? Behav Cogn Psychother [Internet]. 1995 Oct [cited 2020 Jul 16];23(4):325–34. Available from: http://www.cambridge.org/core/journals/behavioural-and-cognitive-psychotherapy/article/what-is-motivational-interviewing/F7E8B9E777291290E6DF0FDE37999C8D

146. Miller WR. Motivational Interviewing with Problem Drinkers. Behav Cogn Psychother [Internet]. 1983 Apr [cited 2020 Jul 16];11(2):147–72. Available from: https://www.cambridge.org/core/journals/behavioural-and-cognitive-psychotherapy/article/motivational-interviewing-with-problem-drinkers/20AD43D18F0976A4DED33EC34FA0C952

147. Miller WR, Rollnick S. Chapter 3—Motivational Interviewing as a Counseling Style [Internet]. Substance Abuse and Mental Health Services Administration (US); 1999 [cited 2017 Oct 7]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK64964/

148. O’Halloran PD, Blackstock F, Shields N, Holland A, Iles R, Kingsley M, et al. Motivational interviewing to increase physical activity in people with chronic health conditions: a systematic review and meta-analysis. Clin Rehabil. 2014 Dec;28(12):1159–71.

149. Hardcastle SJ, Taylor AH, Bailey MP, Harley RA, Hagger MS. Effectiveness of a motivational interviewing intervention on weight loss, physical activity and cardiovascular disease risk factors: a randomised controlled trial with a 12-month post-intervention follow-up. Int J Behav Nutr Phys Act [Internet]. 2013 Mar 28 [cited 2020 Jul 16];10(1):40. Available from: https://doi.org/10.1186/1479-5868-10-40

150. Bandura A. Self-efficacy: The exercise of control. New York, NY, US: W H Freeman/Times Books/ Henry Holt & Co; 1997. ix, 604 p. (Self-efficacy: The exercise of control).

151. Eakin E, Reeves M, Winkler E, Lawler S, Owen N. Maintenance of physical activity and dietary change following a telephone-delivered intervention. Health Psychol Off J Div Health Psychol Am Psychol Assoc. 2010 Nov;29(6):566–73.

152. Ismail K, Bayley A, Twist K, Stewart K, Ridge K, Britneff E, et al. Reducing weight and increasing physical activity in people at high risk of cardiovascular disease: a randomised controlled trial comparing the effectiveness of enhanced motivational interviewing intervention with usual care. Heart [Internet]. 2020 Mar 1 [cited 2020 May 27];106(6):447–54. Available from: https://heart.bmj.com/content/106/6/447

153. Luszczynska A, Schwarzer R. Social cognitive theory. In: Faculty of Health Sciences Publications [Internet]. 2015. p. 225–51. Available from: https://researchbank.acu.edu.au/fhs_pub/5211

154. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: Applications to addictive behaviors. Am Psychol. 1992;47(9):1102–14.

155. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: Toward an integrative model of change. J Consult Clin Psychol. 1983;51(3):390–5.

156. Marshall SJ, Biddle SJ. The transtheoretical model of behavior change: a meta-analysis of applications to physical activity and exercise. Ann Behav Med Publ Soc Behav Med. 2001;23(4):229–46.

157. Korfitsen CB, Larsen RT, Midtgaard J, Christensen J, Juhl CB. Self-Efficacy and Outcome Expectation for Exercise Scales: Cross-cultural Adaptation and psychometric evaluation. 2019;Not published.

158. Resnick B. Reliability and validity of the Outcome Expectations for Exercise Scale-2. J Aging Phys Act. 2005 Oct;13(4):382–94.

159. Resnick B, Jenkins LS. Testing the Reliability and Validity of the Self-Efficacy for Exercise Scale. Nurs Res. 2000;49(3).

160. Lund H, Brunnhuber K, Juhl C, Robinson K, Leenaars M, Dorch BF, et al. Towards evidence based research. BMJ [Internet]. 2016 Oct 21 [cited 2018 Oct 29];i5440. Available from: http://www.bmj.com/lookup/doi/10.1136/bmj.i5440

161. Bandholm T, Christensen R, Thorborg K, Treweek S, Henriksen M. Preparing for what the reporting checklists will not tell you: the PREPARE Trial guide for planning clinical research to avoid research waste. Br J Sports Med [Internet]. 2017 Oct 1 [cited 2020 Feb 5];51(20):1494–501. Available from: https://bjsm.bmj.com/content/51/20/1494

162. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet Lond Engl. 2009 Jul 4;374(9683):86–9.

163. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. BMJ [Internet]. 2010 Mar 24 [cited 2018 Jun 4];340:c869. Available from: https://www.bmj.com/content/340/bmj.c869

164. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med [Internet]. 2009 Jul 21 [cited 2016 Jan 11];6(7):e1000097. Available from: http://dx.doi.org/10.1371/journal.pmed.1000097

165. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 2011 [cited 2016 Apr 11]. Available from: http://handbook.cochrane.org/

166. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013 Feb 5;158(3):200–7.

167. Juhl CB, Lund H. Do we really need another systematic review? Br J Sports Med [Internet]. 2018 Nov 1 [cited 2020 Jul 16];52(22):1408–9. Available from: https://bjsm.bmj.com/content/52/22/1408

168. Garner P, Hopewell S, Chandler J, MacLehose H, Schünemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ [Internet]. 2016 Jul 20 [cited 2018 Jun 27];i3507. Available from: http://www.bmj.com/lookup/doi/10.1136/bmj.i3507

169. Yordanov Y, Dechartres A, Porcher R, Boutron I, Altman DG, Ravaud P. Avoidable waste of research related to inadequate methods in clinical trials. BMJ [Internet]. 2015 Mar 24 [cited 2020 Jul 16];350. Available from: https://www.bmj.com/content/350/bmj.h809

170. Ioannidis JPA. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. Milbank Q. 2016;94(3):485–514.

171. Farrugia P, Petrisor BA, Farrokhyar F, Bhandari M. Practical tips for surgical research: Research questions, hypotheses and objectives. Can J Surg J Can Chir. 2010 Aug;53(4):278–81.

172. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. Trials [Internet]. 2017 Jun 20 [cited 2020 Jul 16];18(3):280. Available from: https://doi.org/10.1186/s13063-017-1978-4

173. Group BMJP. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ [Internet]. 2016 Jul 21 [cited 2018 Oct 16];354:i4086. Available from: https://www.bmj.com/content/354/bmj.i4086

174. Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance the daily amount of physical activity in elderly—a protocol for a systematic review and meta-analysis. Syst Rev [Internet]. 2018 Dec 1 [cited 2018 May 3];7(1):69. Available from: https://link.springer.com/article/10.1186/s13643-018-0733-6

175. Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance amount of physical activity in older adults – a systematic review and meta-analysis. Eur Rev Aging Phys Act [Internet]. 2019 May 4 [cited 2019 May 7];16(1):7. Available from: https://doi.org/10.1186/s11556-019-0213-6

176. Larsen RT, Korfitsen CB, Juhl CB, Boje Andersen H, Langberg H, Christensen J. Concurrent Validity Between Electronically Administered Physical Activity Questionnaires and Objectively Measured Physical Activity in Danish Community-Dwelling Older Adults. J Aging Phys Act [Internet]. 2020 [cited 2020 Dec 12];1–9. Available from: https://journals.humankinetics.com/view/journals/japa/aop/article-10.1123-japa.2020-0214/article-10.1123-japa.2020-0214.xml

177. Larsen RT, Korfitsen CB, Juhl CB, Andersen HB, Christensen J, Langberg H. The MIPAM trial: a 12-week intervention with motivational interviewing and physical activity monitoring to enhance the daily amount of physical activity in community-dwelling older adults – a study protocol for a randomized controlled trial. BMC Geriatr [Internet]. 2020 Dec [cited 2020 Oct 21];20(1):412. Available from: https://bmcgeriatr.biomedcentral.com/articles/10.1186/s12877-020-01815-1

178. Lefebvre C, Manheimer E, Glanville J. Box 6.4.a: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); PubMed format. Cochrane Collaboration; 2011.

179. Higgins J, Savovic J, Sterne JAC, Page M, Hróbjartsson A, Boutron A, et al. A revised tool to assess risk of bias in randomized trials (RoB 2.0). Cochrane Database of Systematic Reviews; 2016.

180. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions: 10.3.1 Fixed-effect method for meta-analysis. 2011.

181. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions: 10.10.4.1 Fixed or random effects? 2011.

182. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions: 10.5.2 Meta-analysis of change scores. 2011.

183. Borenstein M, Hedges L, Higgins J, Rothstein HR, editors. Introduction to meta-analysis. Chichester, U.K: John Wiley & Sons; 2009. 421 p.

184. Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000 Jun;56(2):455–63.

185. Rambøll Management Consulting. SurveyXact. Copenhagen: Denmark; 2020.

186. Moyers TB, Rowell LN, Manuel JK, Ernst D, Houck JM. The Motivational Interviewing Treatment Integrity Code (MITI 4): Rationale, preliminary reliability and validity. J Subst Abuse Treat [Internet]. 2016 Jun [cited 2019 Mar 29];65:36–42. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5539964/

187. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc. 2003 Aug;35(8):1381–95.

188. IPAQ scoring protocol - International Physical Activity Questionnaire [Internet]. [cited 2019 Mar 6]. Available from: https://sites.google.com/site/theipaq/scoring-protocol

189. Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act. 2011 Oct 21;8:115.

190. Mai KS, Sandbaek A, Borch-Johnsen K, Lauritzen T. Are lifestyle changes achieved after participation in a screening programme for Type 2 diabetes? The ADDITION Study, Denmark. Diabet Med J Br Diabet Assoc. 2007 Oct;24(10):1121–8.

191. World Health Organization. Global recommendations on physical activity for health. Genève: WHO; 2010.

192. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy Amst Neth. 1990 Dec;16(3):199–208.

193. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res [Internet]. 2011 Dec [cited 2019 Jul 15];20(10):1727–36. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3220807/

194. Haywood KL, Garratt AM, Fitzpatrick R. Quality of life in older people: a structured review of generic self-assessed health instruments. Qual Life Res Int J Qual Life Asp Treat Care Rehabil. 2005 Sep;14(7):1651–68.

195. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res [Internet]. 2013 [cited 2019 Jul 15];22(7):1717–27. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3764313/

196. Russell DW. UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. J Pers Assess. 1996;66(1):20–40.

197. Lee J, Cagle JG. Validating the 11-Item Revised University of California Los Angeles Scale to Assess Loneliness Among Older Adults: An Evaluation of Factor Structure and Other Measurement Properties. Am J Geriatr Psychiatry [Internet]. 2017 Nov [cited 2019 Aug 8];25(11):1173–83. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1064748117303482

198. Resnick B, Spellbring AM. Understanding What MOTIVATES Older Adults to Exercise. J Gerontol Nurs [Internet]. 2000 Mar 1 [cited 2018 Dec 15];26(3):34–42. Available from: https://www.healio.com/nursing/journals/jgn/2000-3-26-3/%7bf8182f73-42b2-4d8d-9cfb-d494817421e8%7d/understanding-what-motivates-older-adults-to-exercise

199. Resnick B, Orwig D, Wehren L, Zimmerman S, Simpson M, Magaziner J. The Exercise Plus Program for older women post hip fracture: participant perspectives. The Gerontologist. 2005 Aug;45(4):539–44.

200. Due P, Holstein B, Lund R, Modvig J, Avlund K. Social relations: network, support and relational strain. Soc Sci Med 1982. 1999 Mar;48(5):661–73.

201. Lund R, Nielsen LS, Henriksen PW, Schmidt L, Avlund K, Christensen U. Content validity and reliability of the Copenhagen social relations questionnaire. J Aging Health. 2014 Feb;26(1):128–50.

202. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med [Internet]. 2016 Jun [cited 2018 Aug 7];15(2):155–63. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4913118/

203. Peel NM, Paul SK, Cameron ID, Crotty M, Kurrle SE, Gray LC. Promoting Activity in Geriatric Rehabilitation: A Randomized Controlled Trial of Accelerometry. PLOS ONE [Internet]. 2016 Aug 26 [cited 2017 Dec 19];11(8):e0160906. Available from: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0160906

204. Furber S., Butler L., Phongsavan P., Mark A., Bauman A. Randomised controlled trial of a pedometer-based telephone intervention to increase physical activity among cardiac patients not attending cardiac rehabilitation. Patient Educ Couns [Internet]. 2010;80(2):212–8. Available from: http://www.elsevier.com/locate/pateducou

205. Nolan CM, Maddocks M, Canavan JL, Jones SE, Delogu V, Kaliaraju D, et al. Pedometer Step Count Targets during Pulmonary Rehabilitation in Chronic Obstructive Pulmonary Disease. A Randomized Controlled Trial. Am J Respir Crit Care Med. 2017 May 15;195(10):1344–52.

206. Croteau KA, Richeson NE, Farmer BC, Jones DB. Effect of a Pedometer-Based Intervention on Daily Step Counts of Community-Dwelling Older Adults. Res Q Exerc Sport [Internet]. 2007 Dec 1 [cited 2017 Dec 19];78(5):401–6. Available from: http://www.tandfonline.com/doi/abs/10.1080/02701367.2007.10599439

207. Harris T, Kerry SM, Victor CR, Ekelund U, Woodcock A, Iliffe S, et al. A Primary Care Nurse-Delivered Walking Intervention in Older Adults: PACE (Pedometer Accelerometer Consultation Evaluation)-Lift Cluster Randomised Controlled Trial. PLOS Med [Internet]. 2015 Feb 17 [cited 2017 Sep 8];12(2):e1001783. Available from: http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1001783

208. Merom D., Grunseit A., Eramudugolla R., Jefferis B., McNeill J., Anstey K.J. Cognitive benefits of social dancing and walking in old age: The dancing mind randomized controlled trial. Front Aging Neurosci [Internet]. 2016;8(FEB):26. Available from: http://journal.frontiersin.org/article/10.3389/fnagi.2016.00026/full

209. Pelssers J., Delecluse C., Opdenacker J., Kennis E., Van Roie E., Boen F. 'Every step counts!': effects of a structured walking intervention in a community-based senior organization. J Aging Phys Act. 2013;21(2):167–85.

210. Talbot LA., Gaines JM., Huynh TN., Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. J Am Geriatr Soc. 2003;51(3):387–92.

211. Yates T., Davies M., Gorely T., Bull F., Khunti K. Effectiveness of a pragmatic education program designed to promote walking activity in individuals with impaired glucose tolerance: a randomized controlled trial. Diabetes Care. 2009;32(8):1404–10.

212. Lee L-L, Arthur A, Avis M. Evaluating a community-based walking intervention for hypertensive older people in Taiwan: a randomized controlled trial. Prev Med. 2007 Feb;44(2):160–6.

213. Kolbe-Alexander TL, Lambert EV, Harkins JB, Ekelund U. Comparison of two methods of measuring physical activity in South African older adults. J Aging Phys Act. 2006 Jan;14(1):98–114.

214. Deng HB, Macfarlane DJ, Thomas GN, Lao XQ, Jiang CQ, Cheng KK, et al. Reliability and validity of the IPAQ-Chinese: the Guangzhou Biobank Cohort study. Med Sci Sports Exerc. 2008 Feb;40(2):303–7.

215. Rubio Castañeda FJ, Tomás Aznar C, Muro Baquero C. [Validity, Reliability and Associated Factors of the International Physical Activity Questionnaire Adapted to Elderly (IPAQ-E)]. Rev Esp Salud Publica. 2017 Jan 18;91.

216. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. BMC Med Res Methodol [Internet]. 2010 [cited 2016 Oct 19];10:22. Available from: http://dx.doi.org/10.1186/1471-2288-10-22

217. Larsen RT, Wagner V, Keller C, Juhl CB, Langberg H, Christensen J. Feedback from physical activity monitors to enhance amount of physical activity in adults—a protocol for a systematic review and meta-analysis. Syst Rev [Internet]. 2019 Dec [cited 2019 Sep 4];8(1):53. Available from: https://systematicreviewsjournal.biomedcentral.com/articles/10.1186/s13643-019-0970-3

218. Freedland KE, King AC, Ambrosius WT, Mayo-Wilson E, Mohr DC, Czajkowski SM, et al. The Selection of Comparators for Randomized Controlled Trials of Health-Related Behavioral Interventions: Recommendations of an NIH Expert Panel. J Clin Epidemiol [Internet]. 2019 Jun [cited 2020 Jul 18];110:74–81. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6543841/

219. Karlsson P, Bergmark A. Compared with what? An analysis of control-group types in Cochrane and Campbell reviews of psychosocial treatment efficacy with substance use disorders. Addiction. 110(3):420–8.

220. Akinrolie O, Barclay R, Strachan S, Gupta A, Jasper US, Jumbo SU, et al. The effect of motivational interviewing on physical activity level among older adults: a systematic review and meta-analysis. Phys Occup Ther Geriatr [Internet]. 2020 Feb 9 [cited 2020 Jul 1];0(0):1–14. Available from: https://doi.org/10.1080/02703181.2020.1725217

221. Lundahl BW, Kunz C, Brownell C, Tollefson D, Burke BL. A Meta-Analysis of Motivational Interviewing: Twenty-Five Years of Empirical Studies. Res Soc Work Pract [Internet]. 2010 Mar 1 [cited 2018 Dec 16];20(2):137–60. Available from: https://doi.org/10.1177/1049731509347850

222. Magill M, Apodaca TR, Borsari B, Gaume J, Hoadley A, Gordon REF, et al. A Meta-Analysis of Motivational Interviewing Process: Technical, Relational, and Conditional Process Models of Change. J Consult Clin Psychol [Internet]. 2018 Feb [cited 2020 Jul 20];86(2):140–57. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5958907/

223. Cohen-Mansfield J, Marx MS, Guralnik JM. Comparison of exercise models in an elderly population. Aging Clin Exp Res. 2006 Aug;18(4):312–9.

224. Ziebart C, McArthur C, Lee L, Papaioannou A, Laprade J, Cheung AM, et al. “Left to my own devices, I don’t know”: using theory and patient-reported barriers to move from physical activity recommendations to practice. Osteoporos Int [Internet]. 2018 May 1 [cited 2020 Feb 11];29(5):1081–91. Available from: https://doi.org/10.1007/s00198-018-4390-3

225. Shieh C, Weaver MT, Hanna KM, Newsome K, Mogos M. Association of Self-Efficacy and Self-Regulation with Nutrition and Exercise Behaviors in a Community Sample of Adults. J Community Health Nurs. 2015 Oct 2;32(4):199–211.

226. Williams DM, Anderson ES, Winett RA. A review of the outcome expectancy construct in physical activity research. Ann Behav Med Publ Soc Behav Med. 2005 Feb;29(1):70–9.

227. Chen Y, Holahan C, Li X. Loneliness, physical activity, and self-rated health in middle-aged and older adults. Innov Aging [Internet]. 2018 Nov 16 [cited 2020 May 20];2(Suppl 1):967. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6239840/

228. Halaweh H, Willen C, Grimby-Ekman A, Svantesson U. Physical Activity and Health-Related Quality of Life Among Community Dwelling Elderly. J Clin Med Res [Internet]. 2015 Nov [cited 2020 Jul 20];7(11):845–52. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4596265/

229. Brovold T, Skelton DA, Sylliaas H, Mowe M, Bergland A. Association between health-related quality of life, physical fitness, and physical activity in older adults recently discharged from hospital. J Aging Phys Act. 2014 Jul;22(3):405–13.

230. Minocha S, Tudor A-D, Banks D, Holland C, McNulty C, Ail R, et al. Role of digital health wearables in the wellbeing and quality of life of older people and carers [Internet]. The Open University; 2018. Available from:

http://www.niassembly.gov.uk/globalassets/documents/raise/knowledge_exchange/briefing_papers/series7/minocha140318.pdf

231. Seifert A, Schlomann A, Rietz C, Schelling HR. The use of mobile devices for physical activity tracking in older adults' everyday life. Digit Health [Internet]. 2017 Jan [cited 2020 Jul 27];3:205520761774008. Available from: http://journals.sagepub.com/doi/10.1177/2055207617740088

232. Kononova A, Li L, Kamp K, Bowen M, Rikard R, Cotten S, et al. The Use of Wearable Activity Trackers Among Older Adults: Focus Group Study of Tracker Perceptions, Motivators, and Barriers in the Maintenance Stage of Behavior Change. JMIR MHealth UHealth [Internet]. 2019 Apr 5 [cited 2020 Jul 27];7(4). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6473213/

233. Martínez-Alcalá CI, Rosales-Lagarde A, Alonso-Lavernia M de los Á, Ramírez-Salvador JÁ, Jiménez-Rodríguez B, Cepeda-Rebollar RM, et al. Digital Inclusion in Older Adults: A Comparison Between Face-to-Face and Blended Digital Literacy Workshops. Front ICT [Internet]. 2018 [cited 2020 Jul 27];5. Available from: https://www.frontiersin.org/articles/10.3389/fict.2018.00021/full

234. National Academies of Sciences, Engineering, and Medicine, Health and Medicine Division, Board on Population Health and Public Health Practice, Roundtable on Health Literacy. Health Literacy and Older Adults: Reshaping the Landscape: Proceedings of a Workshop [Internet]. Washington (DC): National Academies Press (US); 2018 [cited 2020 Jul 27]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK532464/

235. Batsis JA, Naslund JA, Gill LE, Masutani RK, Agarwal N, Bartels SJ. Use of a Wearable Activity Device in Rural Older Obese Adults: A Pilot Study. Gerontol Geriatr Med [Internet]. 2016 Aug 19 [cited 2017 Sep 8];2:233372141667807. Available from: http://journals.sagepub.com/doi/10.1177/2333721416678076

236. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of Commercially Available Wearable Activity Trackers Among Adults Aged Over 50 and With Chronic Illness: A Mixed-Methods Evaluation. JMIR MHealth UHealth. 2016 Jan 27;4(1):e7.

237. McMahon SK, Lewis B, Oakes M, Guan W, Wyman JF, Rothman AJ. Older Adults' Experiences Using a Commercially Available Monitor to Self-Track Their Physical Activity. JMIR MHealth UHealth [Internet]. 2016 Apr 13 [cited 2017 Aug 24];4(2):e35. Available from: http://mhealth.jmir.org/2016/2/e35/

238. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014 Mar 7;348:g1687.

239. WIJNDAELE K, WESTGATE K, STEPHENS SK, BLAIR SN, BULL FC, CHASTIN SFM, et al. Utilization and Harmonization of Adult Accelerometry Data: Review and Expert Consensus. Med Sci Sports Exerc [Internet]. 2015 Oct [cited 2020 Jul 22];47(10):2129–39. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4731236/

240. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ [Internet]. 1999 Sep 11 [cited 2020 Jul 24];319(7211):670–4. Available from: https://www.bmj.com/content/319/7211/670

241. Morris TP, White IR, Royston P. Tuning multiple imputation by predictive mean matching and local residual draws. BMC Med Res Methodol [Internet]. 2014 Jun 5 [cited 2020 Jul 24];14(1):75. Available from: https://doi.org/10.1186/1471-2288-14-75

242. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol. 2017 Dec 6;17(1):162.

243. Greenland S, Senn SJ, Rothman KJ, Carlin JB, Poole C, Goodman SN, et al. Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. Eur J Epidemiol [Internet]. 2016 Apr [cited 2019 Sep 6];31(4):337–50. Available from: http://link.springer.com/10.1007/s10654-016-0149-3

244. Kraemer HC. Is It Time to Ban the P Value? JAMA Psychiatry [Internet]. 2019 Aug 7 [cited 2019 Sep 6]; Available from: https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2739306

245. Wasserstein RL, Lazar NA. The ASA Statement on p-Values: Context, Process, and Purpose. Am Stat [Internet]. 2016 Apr 2 [cited 2019 Sep 6];70(2):129–33. Available from: https://amstat.tandfonline.com/doi/full/10.1080/00031305.2016.1154108

246. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. J Clin Epidemiol [Internet]. 2014 Mar [cited 2020 Jun 25];67(3):267–77. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3969247/

247. Wickström G, Bendix T. The 'Hawthorne effect' — what did the original Hawthorne studies actually show? Scand J Work Environ Health [Internet]. 2000 [cited 2020 Jul 22];26(4):363–7. Available from: https://www.jstor.org/stable/40967074

248. Barrett S, Begg S, O'Halloran P, Kingsley M. Integrated motivational interviewing and cognitive behaviour therapy can increase physical activity and improve health of adult ambulatory care patients in a regional hospital: the Healthy4U randomised controlled trial. BMC Public Health [Internet]. 2018 Dec [cited 2020 Jun 25];18(1):1166. Available from: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-018-6064-7

249. Holland AE, Mahal A, Hill CJ, Lee AL, Burge AT, Cox NS, et al. Home-based rehabilitation for COPD using minimal resources: a randomised, controlled equivalence trial. Thorax [Internet]. 2017 Jan [cited 2020 May 7];72(1):57–65. Available from: http://thorax.bmj.com/lookup/doi/10.1136/thoraxjnl-2016-208514

250. Savović J, Jones H, Altman D, Harris R, Jűni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. Health Technol Assess Winch Engl. 2012 Sep;16(35):1–82.

251. Rohan Anand, John Norrie, Judy M Bradley, Danny F McAuley, Mike Clarke. Fool's gold? Why blinded trials are not always best. BMJ [Internet]. 2020 Jan 21 [cited 2020 Aug 4];368. Available from: https://www.bmj.com/content/368/bmj.l6228

252. Crutzen R, Göritz AS. Does social desirability compromise self-reports of physical activity in web-based research? Int J Behav Nutr Phys Act [Internet]. 2011 Apr 14 [cited 2020 Jul 22];8:31. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3090315/

253. Durante R, Ainsworth BE. The recall of physical activity: using a cognitive model of the question-answering process. Med Sci Sports Exerc. 1996 Oct;28(10):1282–91.

APPENDIX 1: PAPER A (STUDY I)

Larsen *et al. Systematic Reviews* (2018) 7:69
https://doi.org/10.1186/s13643-018-0733-6

Systematic Reviews

PROTOCOL Open Access

Physical activity monitors to enhance the daily amount of physical activity in elderly—a protocol for a systematic review and meta-analysis

Rasmus Tolstrup Larsen[1*], Jan Christensen[1,2], Carsten Bogh Juhl[4,5], Henning Boje Andersen[3] and Henning Langberg[1]

Abstract

Background: To investigate the use of physical activity monitors (PAMs) for the elderly, the scientific literature should be systematically reviewed and the effect quantified, as the evidence seems inconclusive.

Methods and design: Randomized controlled trials and randomized crossover trials, with participants with a mean age above 65 years, comparing any PAM intervention with other control interventions or no intervention, will be included. This protocol is detailed according to the recommendations of the Cochrane Handbook, and it is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols statement.

Results: We will present results from the search in a flow diagram. The results from the analyses will include regular meta-analyses, stratified analyses, and meta-regressions. The results on each outcome of interest will be presented in a summary of findings table.

Discussion: This paper will explore and analyze the heterogeneity of the results and try to identify variables that will enhance the effect of PAMs in elderly. The results will be useful to researchers working with elderly and/or PAMs, health care professionals working with elderly, and relatives together with the elderly themselves.

Systematic review registration: PROSPERO CRD42018083648.

Keywords: Elderly, Physical activity, Walking, Physical activity monitors, Feedback, Motivation, Behavioral change, Randomized controlled trials, Systematic review, Meta-analysis

Background

Rationale

According to the World Health Organization, elderly who are physically active have lower rates of all-cause mortality and exhibit higher levels of functional health [1]. Even a small change in the daily amount of physical activity may be beneficial on hard outcomes such as all-cause mortality and life expectancy [2]. The American College of Sports Medicine describes walking as the most common type of physical activity among elderly [3] and walking may play a key role in prevention of cardiovascular disease [4].

Physical activity monitors (PAMs) were originally used to quantify the level of physical activity through the amount of steps taken [5] and have been used in research since the 1960s [6]. However, PAMs are also used effectively to motivate and facilitate to an enhanced level of physical activity, as meta-analyses have reported PAM-based interventions to reduce participant weight in weight loss programs [7], increase the level of physical activity [8], and reduce sedentary time significantly [9]. PAMs provide feedback on physical activity [5]. This

* Correspondence: rala@sund.ku.dk
[1]CopenRehab, Section of Social Medicine, Department of Public Health, University of Copenhagen, Copenhagen, Denmark
Full list of author information is available at the end of the article

feedback might facilitate result-driven behavioral change and, hence, increase the level of physical activity [10]. Especially in elderly, PAM-based interventions have been shown to be feasible and effective in enhancing the level of physical activity [11–16]. However, some studies report no significant differences between PAM groups and control groups [17–19]. This could be explained by the devices being difficult and troublesome for elderly to use, and a survey has reported that a third of PAM consumers have stopped using the devices after 6 months [20]. It seems very relevant to investigate the use of PAMs, specifically in elderly as they are expected to behave differently than younger adults. To investigate the above, the scientific literature should be systematically reviewed; the effect quantified and possible factors explaining differences in effect size should be identified.

Research questions and objective

Research questions

1. What is the effect of a PAM-based intervention on physical activity behavior in elderly?
2. What are the potential effects on other outcomes such as changes in body mass index, physical capacity, and health-related quality of life?
3. Which factors explain heterogeneity of the results?

Objective

The objective of this systematic review and meta-analysis is to review the literature and estimate the effect on daily level of physical activity, when using PAMs as an intervention, compared to control interventions in participants aged 65 years and over. Furthermore, to investigate if potential physical activity effects can result in changes in secondary outcomes such as body mass index, physical capacity, and health-related quality of life. Lastly, possible factors explaining heterogeneity will be investigated.

Methods

This protocol is detailed according to the recommendations of the Cochrane Handbook [21], and it is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols (PRISMA-P) statement.

Eligibility criteria

Types of studies

Randomized controlled trials (RCTs) and randomized crossover trials will be included.

Types of participants

Participants included in eligible studies must have a mean age above 65 years. The participants must be independent walkers with or without walking aids.

Types of interventions

For this review, we will include studies comparing any PAM intervention with other control interventions or no intervention. The PAMs may be portable or wearable, electronic or mechanical, and driven by accelerometers, pedometers, or global positioning system (GPS).

Types of outcome measures

Studies must report at least one of the primary or secondary outcomes in order to be included.

Primary outcome The primary outcome is changed in daily amount of physical activity. If more than one relevant outcome is reported within a study, the outcome will be extracted or calculated favoring daily number of steps, followed by daily number of meters walked, daily amount of energy expenditure measured as calories, daily metabolic equivalent of task (minutes or hours), and finally, if no objective measure is available, self-reported physical activity.

Secondary outcomes Secondary outcomes include (a) meeting the study-specific recommended level of physical activity, (b) change in time spent sedentary, (c) change in time spent in moderate activity, (d) change in time spent in vigorous activity, (e) change in physical capacity, (f) changes in body mass index, and (g) changes in self-reported outcomes. All secondary outcomes (a to g) are independent from each other and will be extracted and analyzed accordingly.

a. Meeting the study-specific recommended level of physical activity, measured objectively by PAMs will be extracted from any study that provides results on this.
b. Time spent sedentary will be extracted favoring PAM measured behavior. If no PAM measured

behavior exists, self-reported behavior will be used.

c. Time spent in moderate activity will be extracted favoring PAM measured behavior. If no PAM measured behavior exists, self-reported behavior will be used.
d. Time spent in vigorous activity will be extracted favoring PAM measured behavior. If no PAM measured behavior exists, self-reported behavior will be used.
e. Change in physical capacity will be extracted favoring outcomes measured by a cardiopulmonary exercise test and secondly a walking test such as 6-min walking test or similar.
f. Changes in body mass index will be extracted from any study that provides results on this.
g. Self-reported health-related quality of life will be extracted from any study that provides results on this by using questionnaires. The outcome measurement prioritized in the study will be extracted.

Timing of outcome assessment Data will be extracted at post-intervention and follow-up.

Adverse outcomes All reported adverse events in the eligible studies will be extracted.

Search methods for identification of studies

Electronic searches

Preliminary electronic searches, citation pearl growing, and reference searching have been undertaken to identify relevant references and to screen the papers for relevant search terms. An electronic systematic search for eligible studies in the electronically databases MEDLINE [22], EMBASE [23], SPORTDiscus [24], CINAHL [25], and CENTRAL [26] will be undertaken in March 2018. The search matrix will consist of a combination of relevant keywords and MeSH/Thesaurus terms for (1) population, (2) intervention, and (3) study design. RCTs will be identified using the "The Cochrane highly sensitive search strategies for identifying randomized trials" [27].

No restrictions on language or publication time will be applied. If relevant studies are identified in other language than Danish, English, Swedish, Norwegian, and German, a relevant translator will be contacted. The authors of unobtainable studies or studies with missing data will be contacted.

Search strategy for the systematic search The following search table is used to generate the search strings

Population	Intervention	Study design
"old adults" [Title/Abstract]	"Step counter" [Title/Abstract]	"randomized controlled trial" [Title/Abstract]
elderly[Title/Abstract]	"physical activity monitor" [Title/Abstract]	"controlled clinical trial" [Title/Abstract]
"Above 60 years" [Title/Abstract]	"step monitor" [Title/Abstract]	"cross-over trial" [Title/Abstract]
Seniors [Title/Abstract]	Pedometer [Title/Abstract]	"cross over trial" [Title/Abstract]
Aged [MeSH]	Fitbit [Title/Abstract]	"randomized" [Title/Abstract]
Aged, 80 and over [MeSH]	"activity monitor" [Title/Abstract]	"clinical trial" [Title/Abstract]
Frail elderly [MeSH]	"nokia go" [Title/Abstract]	"randomly" [Title/Abstract]
'older people'. [Title/Abstract]	"misfit ray" [Title/Abstract]	
'older adults' [Title/Abstract]	"moov now" [Title/Abstract]	
	"accelerometer-based tracker" [Title/Abstract]	
	"xiaomi mi band" [Title/Abstract]	
	"tomtom" [Title/Abstract]	
	"vivoactive" [Title/Abstract]	
	"jawbone" [Title/Abstract]	
	"movement counter" [Title/Abstract]	
	"quantified movement" [Title/Abstract]	
	"fitness tracker" [Title/Abstract]	
	"activity monitoring device" [Title/Abstract] physic* AND activit* AND monitor* [Title/Abstract] PAM[Title/Abstract] AND monitor*[Title/Abstract]	

Search string for MEDLINE (((((((("randomly" [Title/Abstract]) OR "clinical trial" [Title/Abstract]) OR "randomized" [Title/Abstract]) OR "cross over trial" [Title/Abstract]) OR "cross-over trial" [Title/Abstract]) OR "controlled clinical trial" [Title/Abstract]) OR "randomized controlled trial" [Title/Abstract])) AND ((((((((((((((((((((PAM[Title/Abstract] AND monitor* [Title/Abstract]))) OR ((physic* AND activit* AND monitor* [Title/Abstract]))) OR "activity monitoring device" [Title/Abstract]) OR "fitness tracker" [Title/Abstract]) OR "quantified movement" [Title/Abstract]) OR "movement counter" [Title/Abstract]) OR "jawbone" [Title/Abstract]) OR "vivoactive" [Title/Abstract]) OR "tomtom" [Title/Abstract]) OR "xiaomi mi band" [Title/

Abstract]) OR "accelerometer-based tracker" [Title/Abstract]) OR "moov now" [Title/Abstract]) OR "misfit ray" [Title/Abstract]) OR "nokia go" [Title/Abstract]) OR "activity monitor" [Title/Abstract]) OR Fitbit [Title/Abstract]) OR Pedometer [Title/Abstract]) OR "step monitor" [Title/Abstract]) OR "physical activity monitor" [Title/Abstract]) OR "Step counter" [Title/Abstract])) AND ((((((((("older people" [Title/Abstract]) OR ("older adults"[Title/Abstract]) OR ("old adults"[Title/Abstract]) OR ("residents"[Title/Abstract]) OR elderly[Title/Abstract]) OR "Above 60 years" [Title/Abstract]) OR Seniors [Title/Abstract]) OR Aged [MeSH]) OR frail elderly [MeSH]) OR ((Aged, 80 and over [MeSH])))

Search string for CINAHL (AB "older adults" OR AB "older people" OR "residents" OR "frail elderly" AB "old adults" OR AB elderly OR AB "Above 60 years" OR AB Seniors OR AB Aged OR AB (Aged, 80 and over)) AND (AB "Step counter" OR AB "physical activity monitor" OR AB "step monitor" OR AB Pedometer OR AB Fitbit OR AB "activity monitor" OR AB "nokia go" OR AB "misfit ray" OR AB "moov now" OR AB "accelerometer-based tracker" OR AB "xiaomi mi band" OR AB "tomtom" OR AB "vivoactive" OR AB "jawbone" OR AB "movement counter" OR AB "quantified movement" OR AB "fitness tracker" OR AB "activity monitoring device" OR AB (physic* AND activit* AND monitor*) OR AB (PAM AND monitor*)) AND (AB "randomized controlled trial" OR AB "controlled clinical trial" OR AB "cross-over trial" OR AB "cross over trial" OR AB "randomized" OR AB "clinical trial" OR AB "randomly")

Search string for EMBASE

1. (old adults or elderly or seniors or residents or older adults or older people or frail elderly).ab.
2. ((PAM and monitor*) or (physic* and activit* and monitor*) or "activity monitoring device" or "fitness tracker" or "quantified movement" or "movement counter" or "jawbone" or "vivoactive" or "tomtom" or "xiaomi mi band" or "accelerometer-based tracker" or "moov now" or "misfit ray" or "nokia go" or "activity monitor" or Fitbit or Pedometer or "step monitor" or "physical activity monitor" or "Step counter").ab.
3. (randomized controlled trial or controlled clinical trial or cross over trial or randomized or clinical trial).ab.
4. aged/
5. controlled clinical trial/ or "randomized controlled trial (topic)"/
6. 1 or 4
7. 3 or 5
8. 2 and 6 and 7

Search string for SPORTDiscus (AB "older adults" OR AB "older people" OR "residents" OR "frail elderly" AB "old adults" OR AB elderly OR AB "Above 60 years" OR AB Seniors OR AB Aged OR AB (Aged, 80 and over)) AND (AB "Step counter" OR AB "physical activity monitor" OR AB "step monitor" OR AB Pedometer OR AB Fitbit OR AB "activity monitor" OR AB "nokia go" OR AB "misfit ray" OR AB "moov now" OR AB "accelerometer-based tracker" OR AB "xiaomi mi band" OR AB "tomtom" OR AB "vivoactive" OR AB "jawbone" OR AB "movement counter" OR AB "quantified movement" OR AB "fitness tracker" OR AB "activity monitoring device" OR AB (physic* AND activit* AND monitor*) OR AB (PAM AND monitor*)) AND (AB "randomized controlled trial" OR AB "controlled clinical trial" OR AB "cross-over trial" OR AB "cross over trial" OR AB "randomized" OR AB "clinical trial" OR AB "randomly")

Search string for CENTRAL

1. Mesh descriptor:[Aged] explode all trees
2. Mesh descriptor:[Frail Elderly] explode all trees
3. "old adults" or elderly or "older adults" or "older people" or "frail elderly" or seniors: ti,ab,kw
4. (PAM and monitor*) or (physic* and activit* and monitor*) or "activity monitoring device" or "fitness tracker" or "quantified movement" or "movement counter" or "jawbone" or "vivoactive" or "tomtom" or "xiaomi mi band" or "accelerometer-based tracker" or "moov now" or "misfit ray" or "nokia go" or "activity monitor" or Fitbit or Pedometer or "step monitor" or "physical activity monitor" or "Step counter":ti,ab,kw
5. (#1 or #2 or #3) and #4

Searching other resources

Searching references of eligible studies and relevant journals by pearl growing will be conducted independently by two reviewers (RTL and JC) in order to include relevant articles not captured by the search strings. The database Clinicatrials.gov will be used to locate ongoing relevant studies.

Data collection and analysis

Data management

The technology platform, *Covidence*, will be used to import citations from the literature searches, screening of title and abstracts, screening of full text, assessing risk of bias in included studies, and extracting the data. The analyses will be conducted in *Stata Statistical Software, version 15*.

Selection of studies
The selection of studies will be done by merging search results from the databases, removing duplicates, examining the titles, and examining full-text reports according to the inclusion criteria.

Two authors (RTL, JC) will independently screen titles and abstracts and assess eligible articles in full text. Any inconsistencies between authors will be discussed and solved with consultation of a third author (CJ).

Data extraction and management
Data on the following items will be extracted from all included studies.

Source Study ID, protocol ID, review author, citation, and contact details

Methods Study design, aim of study, number of arms or groups, funding source, informed consent obtained, and ethical approval

Participants Total number of participants, setting, possible diagnostic criteria, age, sex, country, co-morbidities, education length, and marriage status

Interventions Duration of intervention, specific intervention, and intervention details sufficient for replication

Outcomes All outcomes specified in the "Types of outcome measures" section and specific time points, outcome definitions, and unit of measurement

Results Number of participants allocated to each group, summary data for each intervention, and control groups (as reported) including adverse events

Miscellaneous Funding sources, key conclusions, miscellaneous comments from authors and if correspondence was required

Assessment of risk of bias in included studies
Two review authors (RTL and JC) will independently assess the risk of bias in included studies, using the *RoB 2.0 tool* [28]. Disagreement between review authors will be solved by including a third reviewer (CJ). The risk of bias assessment for each study will be presented using a table with judgment and support for judgment.

Measures of treatment effect
Treatment effect, measured as continuous data, will be expressed as mean difference with 95% confidence intervals for outcomes measured with the same outcome measurement instrument, or as standardized mean difference with 95% confidence intervals, when different measurement instruments are used in the included studies. Dichotomous outcomes, such as adverse events or meeting the recommended level of physical activity, will be analyzed and expressed as risk ratios with 95% confidence intervals.

Unit of analysis issues
Studies with multiple treatment groups If a study has more than one intervention group and both seem relevant for inclusion in the systematic review, the intervention groups will be included as two separate studies and the control group from the study will be separated [29].

Assessment of heterogeneity
The heterogeneity of the study results will be examined using Cochrane Q test and quantified with I^2 values and the between study variance tau^2.

Assessment of small study bias
Small study bias will be assessed by calculating an *Egger's test* score and illustrated with a funnel plot. If small study bias is found, by a positive Egger's test, a *metatrim* analysis will be conducted and an adjusted effect size will be calculated.

Data synthesis
If two or more included studies allows for it, the effect size will be calculated using a random-effects meta-analysis adjusting to Hedges' *g*. If the included studies need network meta-analysis to be pooled this will be performed, as described in Chapter 16.6.3 of the Cochrane Handbook [30]. An alpha level of 0.05 will be considered statistical significant. If it is not possible to conduct a meta-analysis, we will describe the data narratively specific to each outcome.

Subgroup analysis and investigation of heterogeneity
We will explore heterogeneity by conducting sub-group analyses and stratified analyses on the following nominal variables:

- Placement of PAM (ankle worn, wrist worn, hip worn)
- Type of PAM (pedometer versus accelerometer)
- Diagnoses (healthy, cancer patients, pulmonary patients, cardiovascular patients, etc.)
- Feedback frequency (real time, daily, weekly or monthly)

We will explore heterogeneity on continuous data by performing univariate meta-regressions on the following variables:

- Mean age
- Sex distribution
- Number (or percent) of participants with walking aids
- Intervention length
- Mean baseline physical activity

Sensitivity analysis
Sensitivity analyses will be performed on all outcomes of interest by stratifying on overall risk of bias, defined by the RoB 2.0 tool [28] (low/some concerns/high). Furthermore, we will perform sensitivity analyses on how the primary outcome has been extracted, by performing meta-analyses with mean differences on (1) change in daily number of steps, (2) change in daily number of meters walked, (3) change in daily amount of energy expenditure measured as calories, (4) change in daily metabolic equivalent of task (minutes or hours), and (5) self-reported physical activity.

Summary of findings table
We will create a Summary of findings table with effect sizes on the outcomes earlier presented. Two reviewers (RTL and JC) will independently rate the quality of the evidence using the GRADE approach. Downgrading and upgrading the quality of the evidence will be described in the footnotes in the table. An empty example of a summary of findings table is listed in Additional file 1.

Discussion
The aim of this systematic review and meta-analysis is to systematically locate, evaluate, summarize, and analyze available evidence regarding the use of PAMs to enhance the level of physical activity in elderly. This paper will explore and analyze the heterogeneity of the results and try to identify variables that will enhance the effect of PAMs in elderly. The discussion will contain a cost-effective evaluation of the effect size of PAMs and the expected prize of the monitors. We anticipate that this review and the results will be useful to researchers working with elderly and/or PAMs, health care professionals working with elderly, and relatives together with the elderly themselves.

Strength of this systematic review and meta-analysis
We have not identified any systematic reviews investigating the effect of PAM-based interventions in elderly. This paper will be the first systematic review and meta-analysis directly calculating the effect size and exploring heterogeneity of the results. Furthermore, this systematic review and meta-analysis will only include RCTs and randomized crossover trials and hence, the level of evidence is likely to be high.

Additional file

Additional file 1: Table S1. Summary of findings. (DOCX 17 kb)

Abbreviations
GPS: Global positioning system; PAMs: Physical activity monitors; PRISMA-P: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Protocols; RCTs: Randomized controlled trials

Funding
The content presented within this paper was produced as part of the project REACH: this project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 690425 [31].

Authors' contributions
RTL has been the lead author on this protocol and has participated in all steps if the writing process. JC has together with CJ and RTL designed the search strategy and has participated in all steps of the writing process. CJ together with RTL has designed the "Data collection and analysis" section and has participated in all steps of the writing process. HBA and HL together with RTL have participated in the preliminary design of the study and have participated in all steps of the writing process. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
[1]CopenRehab, Section of Social Medicine, Department of Public Health, University of Copenhagen, Copenhagen, Denmark. [2]Department of Occupational and Physiotherapy, Copenhagen University Hospital, Copenhagen, Denmark. [3]DTU Management Engineering Inst., Technical University of Denmark, Diplomvej 372 office 226, 2800 Lyngby, Denmark. [4]Research Unit of Musculoskeletal Function and Physiotherapy, Institute of Sports Science and Clinical Biomechanics, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark. [5]Department of Rehabilitation, Copenhagen University Hospital, Herlev and Gentofte, Denmark.

Received: 3 January 2018 Accepted: 13 April 2018
Published online: 02 May 2018

References
1. World Health Organization. Physical Activity and Older Adults. World Health Organization; 2015. www.who.int/dietphysicalactivity/factsheet_olderadults/en/. Accessed 18 Dec 2017.
2. Wen CP, Wai JPM, Tsai MK, Yang YC, Cheng TYD, Lee M-C, et al. Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. Lancet Lond Engl. 2011;378:1244–53.
3. Chodzko-zajko WJ, Proctor DN, Singh MAF, Minson CT, Nigg CR, Salem GJ, et al. Exercise and physical activity for older adults. Med Sci Sports Exerc. 2009;41:1510–30.
4. Murtagh EM, Murphy MH, Boone-Heinonen J. Walking: the first steps in cardiovascular disease prevention. Curr Opin Cardiol. 2010;25:490–6.
5. Bassett DR, Toth LP, LaMunion SR, Crouter SE. Step counting: a review of measurement considerations and health-related applications. Sports Med. 2017;47:1303–15.
6. Stunkard A. A method of studying physical activity in man. Am J Clin Nutr. 1960;8:595–601.
7. Richardson CR, Newton TL, Abraham JJ, Sen A, Jimbo M, Swartz AM. A meta-analysis of pedometer-based walking interventions and weight loss. Ann Fam Med. 2008;6:69–77.

8. Kang M, Marshall SJ, Barreira TV, Lee J-O. Effect of pedometer-based physical activity interventions: a meta-analysis. Res Q Exerc Sport. 2009;80:648–55.
9. Qiu S, Cai X, Ju C, Sun Z, Yin H, Zügel M, et al. Step counter use and sedentary time in adults. Medicine (Baltimore). 2015;94 https://doi.org/10.1097/MD.0000000000001412.
10. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA. 2015;313:459–60.
11. Harris T, Kerry SM, Victor CR, Ekelund U, Woodcock A, Iliffe S, et al. A primary care nurse-delivered walking intervention in older adults: PACE (pedometer accelerometer consultation evaluation)-lift cluster randomised controlled trial. PLoS Med. 2015;12:e1001783.
12. McMahon SK, Lewis B, Oakes M, Guan W, Wyman JF, Rothman AJ. Older adults' experiences using a commercially available monitor to self-track their physical activity. JMIR MHealth UHealth. 2016;4:e35.
13. Kawagoshi A, Kiyokawa N, Sugawara K, Takahashi H, Sakata S, Satake M, et al. Effects of low-intensity exercise and home-based pulmonary rehabilitation with pedometer feedback on physical activity in elderly patients with chronic obstructive pulmonary disease. Respir Med. 2015;109:364–71.
14. Nishiguchi S, Yamada M, Tanigawa T, Sekiyama K, Kawagoe T, Suzuki M, et al. 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. 2015;63:1355–63.
15. Peel NM, Paul SK, Cameron ID, Crotty M, Kurrle SE, Gray LC. Promoting activity in geriatric rehabilitation: a randomized controlled trial of accelerometry. PLoS One. 2016;11:e0160906.
16. Croteau KA, Richeson NE, Farmer BC, Jones DB. Effect of a pedometer-based intervention on daily step counts of community-dwelling older adults. Res Q Exerc Sport. 2007;78:401–6.
17. Nolan CM, Maddocks M, Canavan JL, Jones SE, Delogu V, Kaliaraju D, et al. Pedometer step count targets during pulmonary rehabilitation in chronic obstructive pulmonary disease. A randomized controlled trial. Am J Respir Crit Care Med. 2017;195:1344–52.
18. Takahashi PY, Quigg SM, Croghan IT, Schroeder DR, Ebbert JO. Effect of pedometer use and goal setting on walking and functional status in overweight adults with multimorbidity: a crossover clinical trial. Clin Interv Aging. 2016;11:1099–106.
19. McMurdo MET, Sugden J, Argo I, Boyle P, Johnston DW, Sniehotta FF, et al. Do pedometers increase physical activity in sedentary older women? A randomized controlled trial. J Am Geriatr Soc. 2010;58:2099–106.
20. Partners E. Inside wearables: how the science of human behavior change offers the secret to long-term…. Endeavour partners. 2017. https://blog.endeavour.partners/inside-wearable-how-the-science-of-human-behavior-change-offers-the-secret-to-long-term-engagement-a15b3c7d4cf3. Accessed 7 Feb 2018.
21. Higgins J, Green S. Cochrane Handbook For Systematic Reviews Of Interventions. 2011. http://handbook.cochrane.org/. Accessed 11 Apr 2016.
22. pubmeddev. MEDLINE database (pubmed). https://www.ncbi.nlm.nih.gov/pubmed. Accessed 3 Oct 2016.
23. Embase database. http://www.embase.com.ez-jmk.statsbiblioteket.dk:2048/#search. Accessed 3 Oct 2016.
24. Inc EIS. SPORTDiscus | Sports Medicine Database | EBSCO. EBSCO Information Services, Inc. | www.ebsco.com. page url. Accessed 21 Nov 2017.
25. EBSCO. CINAHL plus with full text|full text nursing journals | EBSCO | EBSCO health. 2016. https://health.ebsco.com/products/cinahl-plus-with-full-text. Accessed 23 Nov 2015.
26. CENTRAL database. http://onlinelibrary.wiley.com/cochranelibrary/search. Accessed 3 Oct 2016.
27. Lefebvre C, Manheimer E, Glanville J. Box 6.4.a: Cochrane highly sensitive search strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); PubMed format. 2011.
28. Higgins J, Savovic J, Sterne JAC, Page M, Hróbjartsson A, Boutron A, et al. A revised tool to assess risk of bias in randomized trials (RoB 2.0). 2016. https://sites.google.com/site/riskofbiastool/20161020_RoB_2.0_template_parallel_assignment.docx?attredirects=0&d=1.
29. Higgins, Green. 16.5.4 How to include multiple groups from one study. 2011. http://handbook-5-1.cochrane.org/chapter_16/16_5_4_how_to_include_multiple_groups_from_one_study.htm. Accessed 4 Dec 2017.
30. Higgins JPT, Green S. 16.6.3 Multiple-treatments meta-analysis—Cochrane Handbook for Systematic Reviews Of Interventions. 2011. http://handbook-5-1.cochrane.org/. Accessed 19 Oct 2017.
31. REACH | HORIZON 2020. http://reach2020.eu/. Accessed 23 Nov 2017.

APPENDIX 2: PAPER B (STUDY I)

Larsen *et al. European Review of Aging and Physical Activity* (2019) 16:7
https://doi.org/10.1186/s11556-019-0213-6

European Review of Aging and Physical Activity

REVIEW ARTICLE Open Access

Physical activity monitors to enhance amount of physical activity in older adults – a systematic review and meta-analysis

Rasmus Tolstrup Larsen[1*], Jan Christensen[2,6], Carsten Bogh Juhl[4,5], Henning Boje Andersen[3] and Henning Langberg[1]

Abstract

Background: The body of evidence related to the effect of physical activity monitor-based interventions has grown over the recent years. However, the effect of physical activity monitor-based interventions in older adults remains unclear and should be systematically reviewed.

Objective: The objective of this systematic review was to estimate the effect of physical activity monitor-based interventions on physical activity behavior in participants aged 65 and above. Subsequently we explored the effect on body mass index, physical capacity, and health-related quality of life and finally the impact of patient- and intervention characteristics.

Methods: Searches in MEDLINE, EMBASE, SPORTDiscus, CINAHL, and CENTRAL were performed on April 26, 2018. No publication date filters were applied. References of eligible studies were scrutinized and relevant journals were hand-searched. Randomized controlled trials and randomized cross-over trials investigating the effect of a physical activity monitor-based intervention on physical activity were included. Studies were included if the mean age of the participants was above 65 years, and participants could walk independently with or without walking aids. The Cochrane handbook was used as a template for extracting data and the RoB 2.0 tool was used to assess risk of bias. Random-effects meta-analysis using Hedges g, were used to pool the study results. The main outcome of this study was physical activity.

Results: Twenty-one studies with 2783 participants were included. The median participant age in the studies was 70.5 years, the median percentage of male participants was 42%, and the median baseline daily step count was 5268. Physical activity monitor-based interventions had a moderate effect (SMD = 0.54, 95% CI: 0.34 to 0.73) compared to control interventions, corresponding to an average increase of 1297 steps per day in the intervention groups. No impact of patient and intervention characteristics on the effect estimates were found.

Short conclusion: Low quality of evidence was found for a moderate effect of physical activity monitor-based interventions on physical activity compared with control interventions. More studies with higher research methodology standards are required.

PROSPERO registration: CRD42018083648.

Keywords: Aging, Physical activity, Older adults, Physical activity monitors, Technology, Motivation, Systematic review, Meta-analysis, Walking, Moderate to vigorous physical activity

* Correspondence: rala@sund.ku.dk
[1]CopenRehab, Department of Public Health, Section of Social Medicine, University of Copenhagen, Gothersgade 160, 3rd floor, 1123 Copenhagen K, Denmark
Full list of author information is available at the end of the article

Background

Physical inactivity is a growing worldwide problem and it has been reported to cause 9% of all premature death [1]. The amount of daily physical activity (PA) decreases with age [2–5] and one in eight European adults age 55 or older never or hardly ever, engage in moderate to vigorous PA (MVPA) [6]. Functional decline is expected and unavoidable in older adults, but regular exercise can minimize the physiological effects of an otherwise sedentary lifestyle and thus increase life expectancy by improving function of daily living and by slowing progression of disease and disability [7].

An older systematic review reported that physical activity monitor (PAM)-based interventions significantly enhanced the amount of PA with an average of 2491 steps per day, compared to the control group interventions, in adults [8]. Among older adults, the use of PAMs has been reported to be feasible [9, 10] and several recently published randomized controlled trials (RCTs) report promising results [11–16]. However, these studies differ with respect to sample characteristics, intervention length and setting, which might have resulted in the differences in the reported effect sizes between studies [11–16].

The body of evidence related to the effect of PAM-based interventions has grown over the recent years. However, the effect of PAM-based interventions in older adults remains unclear and should be systematically reviewed. Further, patient and intervention characteristics should be explored to understand their impact on PA levels. This information may be used to inform future research and provide guidance to clinical decision-makers considering the use PAMs in PA programs.

Objective

The objective of this systematic review and meta-analysis was to estimate the effect of PAM-based interventions on amount of PA (e.g. daily step count) in participants aged 65 and above. Subsequently we aimed to explore the effect on time spent sedentary, MVPA time, physical capacity (e.g. measured by a cardiopulmonary exercise test or as a walking test), body mass index (BMI), and health-related quality of life (HRQoL) (e.g. by questionnaires). Finally, we sought to investigate the impact of participant- (e.g. diagnoses, age and sex distribution), intervention- (e.g. intervention length, type of PA measure and feedback frequency) and study (e.g. risk of bias) characteristics on the results.

Methods

Protocol and registration

This systematic review and meta-analysis is detailed according to the recommendations of the Cochrane Handbook [17] and it is reported according to the PRISMA statement. The method is described in the published review protocol as well as in the PROSPERO registration (CRD42018083648) [18, 19]. Unless otherwise stated, the methods used and reported in this systematic review followed the review protocol.

Eligibility criteria

We included RCTs and randomized cross-over trials comparing any PAM-based intervention where the participants of the intervention group received any kind of feedback on their physical activity level measured by PAMs, and where the control intervention did not receive feedback from the PAMs. The mean age of the participants should be above 65 years, and participants should be able to walk independently with or without walking aids.

The primary outcome was PA. If more than one type of PA measure were reported, we extracted or calculated it in the following order: daily number of steps, daily number of meters walked, daily amount of energy expenditure (calories), daily metabolic equivalent of task (minutes or hours) and finally, if no objective measure was available, self-reported PA. The secondary outcomes included:

- Time spent as sedentary (measured objectively by PAMs)
- Time spent in MVPA (measured objectively by PAMs or secondly as self-reported behavior)
- Physical capacity (measured by a cardiopulmonary exercise test or secondly as a walking test)
- BMI
- Self- reported HRQoL determined by questionnaires.

End-point scores were used to calculate treatment effects. To avoid unit-of-analysis error with cross-over trials, the outcome was extracted at baseline and when the first period ended, as recommended in the Cochrane Handbook chapter 16.4.5 [20]. Reported adverse events or withdrawals due to illness were extracted if possible.

Information sources

Preliminary searches and identification of relevant papers were performed to identify relevant search terms and subject headings. The final systematic search for eligible studies in MEDLINE, EMBASE, SPORTDiscus, CINAHL, and CENTRAL was performed on April 26, 2018. Additional studies that met the inclusion criteria were obtained through an independent review of article references by two reviewers (RTL and JC). The Clinicatrials.gov database was searched on February 13th 2018 to locate ongoing relevant studies.

Search

The search string consisted of a combination of relevant keywords and subject headings for: PAMs, older adults, and randomized studies and can be found in the study protocol [19]. No restrictions on language or publication-time were applied. The authors of unobtainable studies or studies with missing data were contacted to obtain missing information.

Study selection, data items and data collection process

Citations was imported into the technology platform, *Covidence.* Two authors (RTL and JC) screened the titles and abstracts independently and assessed full-text reports. Any inconsistencies between authors was discussed and rectified in consultation with a third author (CJ). Data extraction was performed independently by two authors (RTL and JC).

Risk of bias in individual studies

Two review authors (RTL and JC) independently assessed study quality using the Risk of Bias 2.0 tool [21] on study outcome level. Disagreement was solved by discussion with a third reviewer (CJ).

Summary measures

Treatment effects, on continuous data, were expressed as standardized mean differences (SMD) with 95% confidence intervals. The SMD was translated back to a mean difference in steps for the primary outcome, MVPA time, meters on 6MWT for physical capacity and BMI (kg/m^2) respectively, using the method described by Bliddal and Christensen [22]. The SDs used for translating the SMDs were extracted for each outcome from the intervention group from largest study with the lowest risk of bias, in which objectively measured values were favored. The SDs used were 2402 steps per day [23], 16.2 min of daily MVPA [14], 80 m on a 6MWT [24] and 4.8 kg/m^2 on BMI [25]. When reported in text or study flow diagrams, adverse event and participant withdrawal rates were extracted and expressed as relative risks with 95% confidence intervals. If a study reported zero adverse events, the Der-Simonian & Laird method was used and 0.5 was added as a value to enable random effects meta-analysis [26].

Synthesis of results

The effect size was calculated using a random-effects model adjusting to Hedges' *g*, using end-point scores only. In studies where no continuous data were available for the outcomes, we used dichotomous data and converted the odds ratios and the standard errors (log ES) into the standard mean difference using the Chinn et al. approach described in chapter 9.4.6 of the Cochrane Handbook [27]. An alpha level of 0.05 was considered statistically significant. Stata/IC 15.1 for Mac (64-bit Intel), Copyright 1985–2017 StataCorp LLC was used for all statistical analyses.

Unit of analysis issues

One study had two relevant intervention groups [11]. The intervention groups of the study were included as two separate comparisons and the control group from the study was separated according to guideline in the Cochrane Handbook chapter 16.5.4 [28].

Additional analyses

The heterogeneity of the extracted results was examined using the Cochrane Q test and quantified with I^2 statistic. We performed subgroup analyses to explore the impact of characteristics of participants and intervention and stratified the effect size on the following nominal variables: type of PAM (accelerometer versus pedometer), diagnoses of participants (none, cardiac patients, COPD or osteoarthritis), feedback frequency (daily, weekly or monthly) and overall risk of bias (low, some concerns and high). We conducted three explorative subgroup analyses: one analysis on control intervention content (advice group, goal setting, maintain usual PA, other training, rehabilitation program, and usual care), one analysis on grouping the interventions into types (gamification, incremental goals, monthly feedback and reinforcement, ongoing counseling and pre-counseling), and one analysis on active control interventions versus non-active control interventions (maintain usual PA or no intervention).

We chose to investigate how the method of physical activity reporting (i.e., objective measurement, self-report, interview) affected the results. This sensitivity analysis was deemed more informative than the protocolled sensitivity analysis on mean differences in daily number of steps, daily number of meters walked, daily amount of energy expenditure measured as calories, daily metabolic equivalent of task, and self-reported physical activity.

Publication bias were assessed by Eggers test. If small study bias was present, the Duval and Tweedie nonparametric "trim and fill" analysis was conducted adjusting the effect size [29, 30].

We performed univariate meta-regressions on continuous data on the following variables:

Age (years), sex distribution (percent), number (or percent) of participants with walking aids, intervention length (weeks), baseline PA (steps), and BMI.

Results

Study selection

Twenty-one studies were included in the review [11–16, 23–25, 31–42]. We identified one ongoing trial (Clinicaltrials.gov Identifier: NCT03086850), but we did not include this study as it was in the participant recruitment phase. Citations and reasons for exclusion from full text screening are listed in the Additional file 1: Table S3. The study selection process is illustrated in Fig. 1. A summary of the included studies is listed in Table 1. Characteristics of the 21 included studies (22 comparisons, 2783 participants) are listed in Additional file 1: Table S2.

Study characteristics

Risk of bias within studies

The risk of bias summary and review authors' judgements about each risk of bias item are presented in Fig. 2. Figure 3 illustrates the risk of bias as percentages across all included studies for each risk of bias item. Overall, five studies were considered as having a low risk of bias [13, 14, 23, 25, 40], 10 studies were considered as having some concerns [12, 15, 16, 24, 31, 34–36, 38, 39], and six studies were considered as having a high risk of bias [11, 12, 32, 33, 37, 41]. Judgements and support for judgement about each item is presented for all studies in characteristics of studies in Additional file 1.

In two studies the risk of bias assessment differed between outcomes. Kolt et al. and Nolan et al. were assessed to have high risk of selective outcome reporting, on self-reported HRQoL (SF-36) [25, 31].

Synthesis of the results and effect of the interventions

Twenty studies (21 study comparisons and 2704 participants) evaluated the effect of PAM on PA. The random effects meta-analysis is illustrated in Fig. 4. The overall SMD was 0.54, (95% CI: 0.34 to 0.73), $I^2 = 79.2\%$, $p < 0.001$, favoring the PAM interventions. When using a SD of 2402 steps per day, this corresponds to a weighted mean difference of 1297 (95% CI: 817 to 1753) favoring the intervention groups [23].

Secondary outcomes

Only one study (35 participants) reported the effect of the intervention on time spent sedentary [13]. The SMD of this study was calculated to be − 0.40 (95% CI: -1.07 to 0.27), favoring the PAM intervention. The difference in weekly sedentary time was 44.0 min (95% CI: 37.1 to 50.9) with the control group being most sedentary.

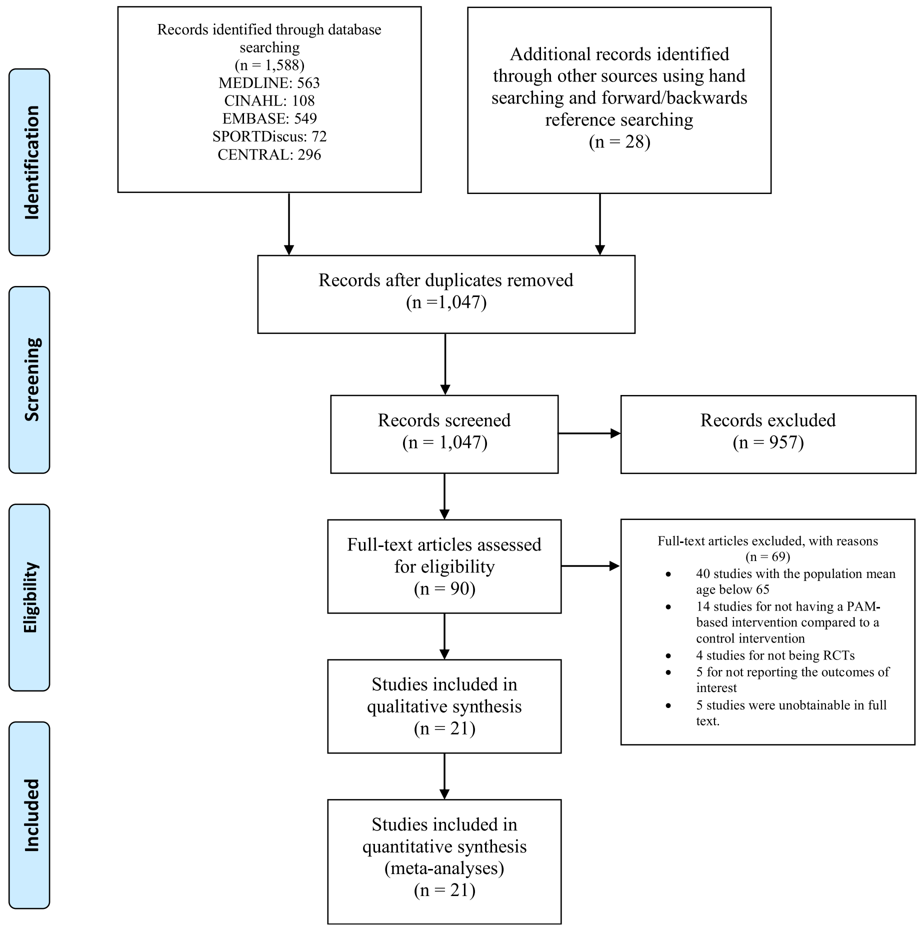


Fig. 1 PRISMA flow diagram illustrating the inclusion process

Table 1 Summary of the characteristics of the included studies. Citations of studies that reported results on domains are listed after the domain

Methods	Number of studies (%)
RCT with parallel group design [11–16, 23–25, 31, 32, 34–42]	20 (95%)
RCT with cross over design [33]	1 (5%)
Setting	Number of studies (%)
Europe [12, 13, 23, 24, 31, 32, 35, 38]	8 (38%)
Australia and New Zealand [14, 16, 25, 37]	4 (19%)
Asia [15, 36, 40, 41]	4 (19%)
North America [11, 33, 34, 39, 42]	5 (24%)
Participant diagnoses	Number of studies (%)
Osteoarthritis [23]	1 (5%)
COPD [12, 31, 36]	3 (14%)
Cardiac patients [16, 40]	2 (10%)
None [11, 13–15, 24, 25, 32–35, 37–39, 41, 42]	15 (71%)
Participant characteristics	Median (range)
Median age in studies (k = 21)	70.5 (65 to 81.5)
Median body mass index in studies [11, 12, 14, 15, 23, 25, 31, 35, 36, 39, 40, 42]	27.9 (21.1 to 31.82)
Median percentage of male participants in studies [11–16, 23–25, 31, 33–36, 38–42]	42 (0 to 88)
Median percentage of married participants [13, 16, 24, 25, 32, 35, 39]	61.4 (39 to 80.5)
Median baseline daily step count [11–13, 15, 23, 31, 33–35, 39, 41]	5268 (2420 to 7697)
Intervention	Median (range)
Length median weeks (k = 21)	12 (4 to 52)
Physical activity monitor	Number of studies (%)
Accelerometer [12, 14, 32, 36, 41]	5 (24%)
Pedometer [11, 13, 15, 16, 23–25, 31, 33–35, 37–40, 42]	16 (76%)
Frequency of feedback	Number of studies (%)
Daily [11–13, 15, 16, 23–25, 32, 33, 35, 37–42]	17 (81%)
Weekly [14, 31, 33]	3 (14%)
Monthly [36]	1 (5%)
Outcomes	Number of studies (%)
Reported results on physical activity [11–16, 23–25, 31–36, 38–42]	20 (95%)
Reported results on sedentary time [13]	1 (5%)
Reported results on MVPA time [14, 24, 25, 31, 35, 41, 42]	7 (33%)
Reported results on physical capacity [24, 36, 37, 41]	4 (19%)
Reported results on health-related quality of life [13, 24, 25, 31, 36]	5 (24%)
Reported results on body mass index [25, 36]	3 (14%)
Reported results on adverse events [13, 14, 16, 24, 31, 32, 35–37, 40]	10 (48%)

RCT Randomized Controlled Trial, *COPD* Chronic Obstructive Pulmonary Disease, *MVPA* Moderate to Vigorous Physical Activity, *k* number of studies. The reported median of mean values are unweighted in relation to study size or reporting precision

A total of eight studies (1686 participants) reported data on effect of the interventions on MVPA time. The overall SMD was 0.34 (95% CI: 0.15 to 0.52), $I^2 = 65.8\%$, $p = 0.005$, favoring the PAM interventions. When using a SD of 16.2 of daily MVPA, this corresponds to a weighted mean difference of 5.5 min per day (95% CI: 2.4 to 8.4) with more MVPA in the intervention groups [14].

A total of four studies (754 participants) reported the effect of the intervention on physical capacity. The overall SMD was 0.19 (95% CI: -0.10 to 0.48), $I^2 = 48.8\%$, $p = 0.118$, favoring the PAM intervention. When using a SD of 80 m, this corresponds to a weighted mean difference on 15 m (95% CI: -8 to 38) with more meters walked on a 6MWT in the intervention groups [24].

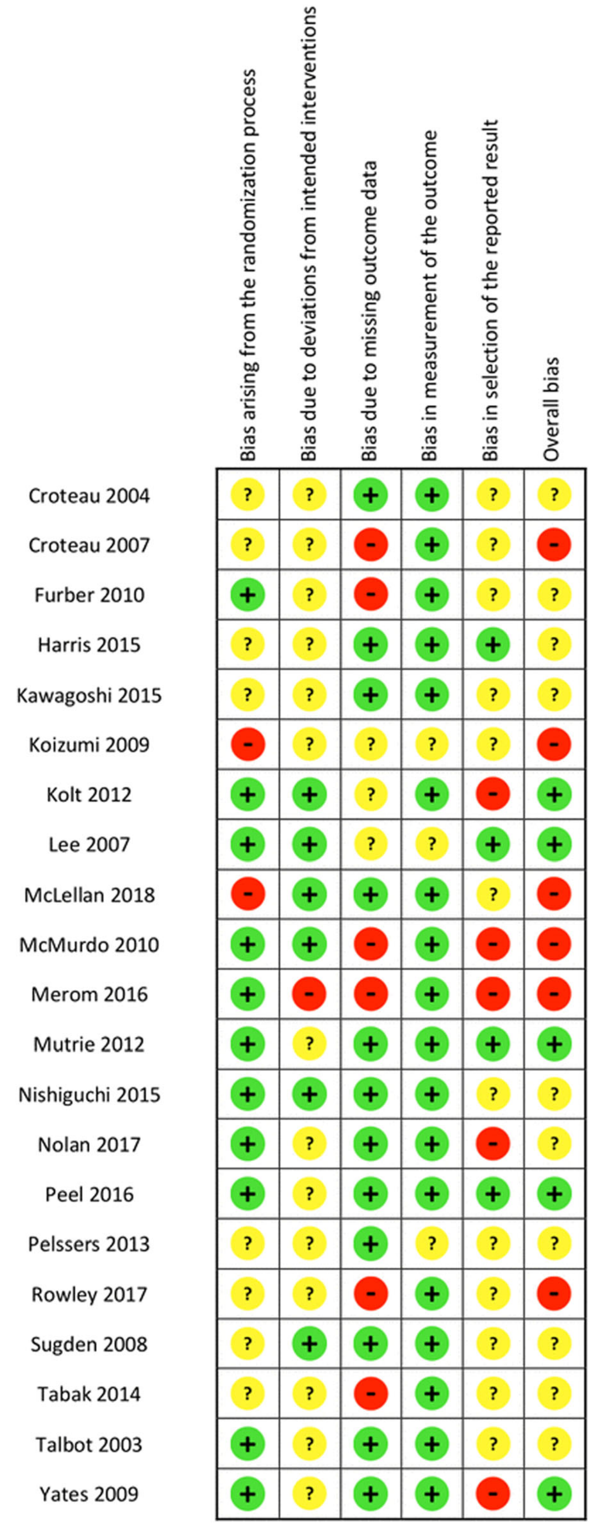


Fig. 2 Risk of bias summary: review authors' judgements about each risk of bias item for each included study. +: Low risk of bias,?:Some concerns, %: High risk of bias

A total of three studies (570 participants) reported data for effect of the interventions on BMI. The overall SMD was 0.15, (95% CI: -0.01 to 0.31), $I^2 = 0\%$, $p = 0.752$, favoring the control intervention. When using a SD of 4.8 kg/m^2, this corresponds to a mean difference on 0.72 kg/m^2 (95% CI: -0.05 to 1.50) with the control groups having the lowest BMI [25].

A total of five studies (1038 participants) reported data for effect of the interventions on HRQoL. The overall SMD was 0.01, (95% CI: -0.12 to 0.14), $I^2 = 0.0\%$, $p = 0.541$, favoring the PAM interventions.

A summary of the analyses on the secondary outcomes are illustrated in Fig. 5.

Meeting the study specific recommended level of physical activity

No studies reported data on this.

Additional analyses

Subgroup analyses on the type of PAM, diagnoses, feedback frequency, risk of bias judgement and type of PA measure on the effect of the intervention on PA (Additional file 1: Figure S6), MVPA time (Additional file 1: Figure S7), physical capacity (Additional file 1: Figure S8), BMI (Additional file 1: Figure S9), and HRQoL (Additional file 1: Figure S10) are presented in the Additional file 1. No significant differences in the subgroup analyses were observed for any outcomes.

Additional file 1: Table S3 reports data from sensitivity-analyses (univariate meta-regressions) on how the SMD from all outcomes sere affected by the following variables: age in years, sex distribution in percent male, percent of participants with walking aids, intervention length in weeks, baseline PA measured in steps per day, BMI in kg/m^2. None of the above-mentioned variables were significantly correlated with the effect size for any outcomes, nor did any variable reduce Tau^2 statistic. There were insufficient observations to analyze the correlation between effect size and percent of participants with walking aids for all outcomes.

Egger's test showed significant small study bias for the effect on PA ($p = 0.036$), indicating that the analyzes are overestimating the effect on PA (Additional file 1: Figure S11). The bias adjusted (trimmed and filled) analysis with random effects revealed an adjusted SMD on 0.37, (95% CI: 0.15 to 0.59) after filling the analysis with three fictive studies. Analyzing the effect on time spent in MVPA, physical capacity, BMI, and HRQoL, no small study bias was found using Egger's test.

Adverse events

A total of 11 studies (1927 participants) reported data for adverse events. The overall relative risk for adverse events was 0.91, (95% CI: 0.66 to 1.25), $I^2 = 0.0\%$ $p = 0.942$, with

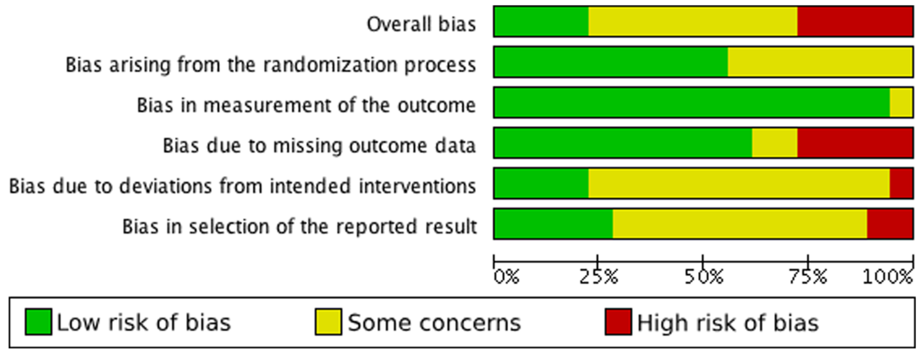


Fig. 3 Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies

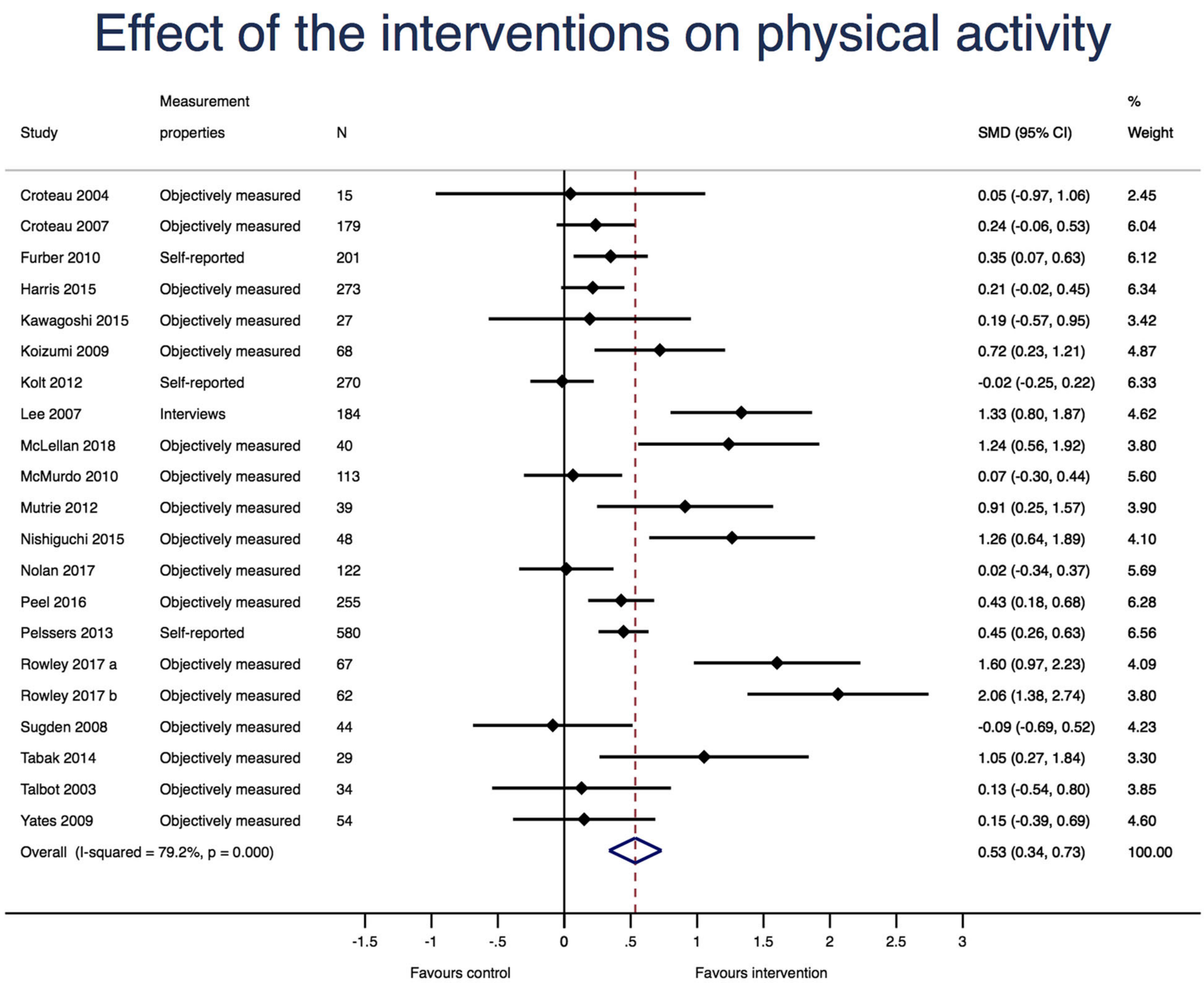


Fig. 4 Random effects meta-analysis with effect of the interventions on physical activity using Hedges g. N: Number of participants; SMD; standardized mean difference. For each study, the diamond represents the standardized mean difference of the intervention effect with the horizontal line representing 95% confidence intervals. The large diamonds represent the pooled standardized mean difference between the intervention groups and the control groups

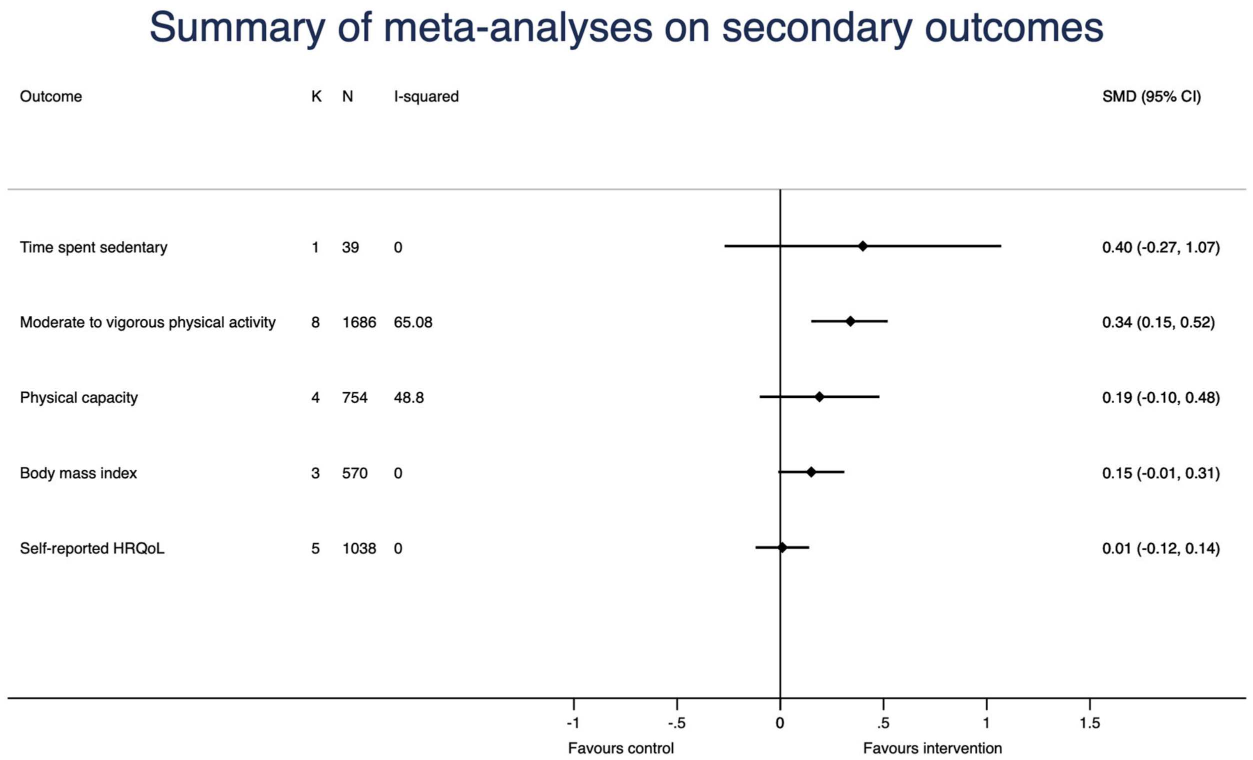


Fig. 5 Summary of random effects meta-analyses with effect of the interventions on secondary outcomes. K: number of studies; N: Number of participants; SMD: standardized mean difference; HRQoL: health-related quality of life. For each analysis, the diamond represents the standardized mean difference of the pooled intervention effect with the horizontal line representing 95% confidence intervals

more adverse events and withdrawals due to illness in the control groups. The random effects meta-analysis for the adverse events is shown in Additional file 1: Figure S12.

Explorative post-hoc subgroup analyses

Additional file 1: Figure S13, illustrates an explorative analysis of effect of interventions on PA sorted on type of control intervention and Additional file 1: Figure S14, illustrates an explorative analysis of effect of interventions on PA sorted on groupings of intervention types. However, none of the findings were significant. Additional file 1: Figure S15, illustrates an explorative analysis of effect of interventions on PA sorted on active control interventions versus non-active control interventions. The 11 study comparisons (1219 participants) with non-active control interventions had a significantly larger effect compared to the 10 study comparisons (1485 participants) with an active control intervention.

Discussion

The objective of this systematic review was to investigate the effect of PAM-based interventions on older adults.

Our primary outcome of interest was PA and the main results include a moderate effect, equivalent to a larger increase on 1297 more steps per day in the intervention groups and the small to moderate effect on MVPA time equivalent to a larger increase on 8 more minutes per day in the intervention groups. As we were not able to explain the heterogeneity of the results with any of our sub- or sensitivity analyses, the effect of the interventions may be applicable to the broadly defined older adult population. However, further potential influences, such as medication and disease specific treatments need to be considered.

In terms of translating the effect on 1297 more steps per day, there is a lack of evidence on how much is clinically relevant change in general, for older adults. The WHO recommends that older adults are equally physically active as their younger counterparts but if co-morbidities limits their ability to be physically active, they should be as active as their conditions allow [43]. A systematic review suggests that the WHO-recommended 30 min of MVPA per day is equivalent to 7000 to 10,000 steps per day in older adults [43, 44]. According to Table 1, the median baseline daily step count in the studies was 5268 which makes the effect on 1297 steps equivalent to a 25% increase in daily number of steps. If the effect size is added to the median baseline daily step count, the average older adult will get close to 7000 steps per day. This highlights the clinical relevance of the results. Other more invasive exercise interventions may be more effective in increasing the amount of daily PA in older adults, but as PAM based interventions are not very invasive, they could be implemented in large scale projects as well.

This review provides evidence for the use of PAMs as an intervention to promote PA among older adults. Our finding of a moderate effect is in line with a former systematic review by Bravata et al. that estimated the effect size to be 2491 steps per day (95% CI: 1098 to 3885) in a population with a mean age on 49 years [8]. The population of interest in the Bravata et al. systematic review is more than 20 years younger than the median mean age in the included studies from this review [8]. As the level of physical activity decrease with age [2–5], a younger population is expected to be more active which may explain why the effect in steps per day is almost twice as large in the Bravata et al. systematic review [8]. However, as the effect sizes are not significantly different from each other, the above-mentioned explanation is only relevant if future systematic reviews find a significant effect modification from age, which we did not find in this review.

Even though we only included one study with results on sedentary time [13], this study was also included in a recent published systematic review from Qui et al. that reports PAM usage to be significantly associated with reduced sedentary time among adults [45].

Among older adults, level of PA is associated with, age, BMI and sex [46]. Contrary to this we were not able to explain the variance in the effect of the interventions through participant age, BMI or sex. However, this also means that we did not find any specific subgroup of older adults that may not benefit from using PAMs to enhance the level of physical activity.

The prevalence of frailty and chronic diseases are high in older adults [47, 48]. At first glance, our results could be limited to older adults with a higher function and a lower disease prevalence as the majority of the included studies included community dwelling older adults without specific diseases [11, 13–15, 24, 25, 32–35, 37, 38, 41, 42]. However, among these studies, several samples were inactive or did not meeting PA recommendations [11, 13, 32, 38]. One study was conducted in a post-acute care rehabilitation setting [14] and other studies included patients with hypertension [24], osteoarthritis [34], cancer [34], and other chronic diseases [35]. Four studies did not describe the disease characteristic of the participants [15, 24, 41, 42]. The broad range of participant characteristics across studies is a strength of this systematic review as it increases the generalizability of the findings to the general population.

None of the subgroup analysis showed any significant impact of risk of bias on the effect. We did however find an overestimation of the effect size on the PA caused by small study bias. Publication bias will normally overestimate the effect of the published interventions due to type 1 errors or selective outcome reporting [49]. In summary, we have chosen to downgrade the overall quality of the evidence due to publication bias.

We conducted three additional analyses to investigate the impact of intervention and control intervention content. PAM-based interventions had a significant greater effect in studies with non-active control interventions compared to studies with active control interventions. No other effects were significant. This analysis is recommended to obtain a meaningful estimate of the effect of the interventions and to avoid a confused picture of absolute intervention effects [50]. Using non-active controls will by nature give a larger effect size, as most interventions (also control interventions) will have some effect. Thus, future studies should use direct comparisons to investigate if PAMs can be an effective add-on *intervention, or if other types of behavior change strategies can effectively increase the effect from the PAMs.*

A Hawthorne-effect, meaning that the participants in control groups could be expected to increase their level of PA, simply due to participation in a PA study may be present in the included studies. This was also discussed in the systematic review published by Bravata et al. [8]. When comparing a PAM-based intervention where the participants receive feedback to a control group that are aware that they are being measured, the effect size might be slightly underestimated when compared to PAM-based walking programs with no control groups. This has been addressed systematically by Waters et al., who found a similar effect in both control and intervention groups in eight of 29 PA trials [51]. This is in line with our explorative results from Additional file 1: Figure S13 that illustrates a larger effect in studies that includes control groups who were asked to maintain usual PA and Additional file 1: Figure S15 that illustrates a larger effect size in studies that uses non-active control interventions. This may be explained by participants who volunteer for trials because they wish to increase their level of PA, participants being refractory or nonadherent after being allocated to a control group and several other factors which should be kept in mind when interpreting results from PA trials or reviews.

Limitations

There were some deviations from the published study protocol [19]. Firstly, there were insufficient data to determine if participants met the study-specific recommendation for level of physical activity. We proposed to study this outcome in our protocol; however, none of the studies included in this review reported on this outcome. Secondly, we chose to pool the moderate and vigorous activity as most of the included studies did not distinguish between these intensity categories in their reporting.

We performed a wide and comprehensive literature search across several relevant databases, and used a pearl growing strategy where two reviewers independently located relevant references through journal sites and reference lists of included studies. Additionally, we obtained relevant references by using forward and backwards reference searches. Despite this wide and robust search strategy, it is possible that not all relevant studies were included in this systematic review.

In terms of translating the SMDs back to number of steps, MVPA time, meters walked in a 6MWT and BMI, the translation should only be read as a way of making our results easier to interpret and comes with limitations to generalizability. Firstly, we assume that the SMD can be used to extrapolate results, but some studies used different scales and outcome measures which might bring some problems. Secondly, the true SD of the population is impossible to estimate. However, when choosing the SDs from the largest study with the lowest risk of bias rating we have tried to be transparent and avoid bias in the selection. It should be noted that interpretation must happen with caution as it basically only represents the study from which the SD was chosen.

This systematic review is focused on older adults above the age of 65 years. As reported in Table 1, some of the studies will include results from participants younger than 65 which might bring some bias to our external validity. However, according to Additional file 1: Table S3, the association between study mean age and the effect size was clearly not significant for all outcomes and the study mean age explained almost no effect size heterogeneity. We hereby acknowledge the limitation that some included studies would have had younger participants, but we find no evidence for affecting the external validity to the population of interest.

Body of evidence

We found that the quality of the body of evidence of PAM-based interventions was low to moderate. Our results were affected by unexplained heterogeneity, publication bias and imprecision. The pooled effects for time spent sedentary, physical capacity, BMI and self-reported HRQoL were not significant. Furthermore, the confidence interval for the effect size of the primary outcome, PA, suggests that the overall effect is small to moderate. However, a moderate quality of evidence was found on the risk of adverse events being the same in the intervention and the control groups. PAMs seem useful for public health interventions as it seems to be safe and effective to include them in PA programs for old adults. The grading of the body of evidence for each outcome is reported in the summary of findings table (Additional file 2).

Conclusion

General interpretation of results

This review demonstrates low quality of evidence for a moderate effect on PA, equivalent to a larger increase at 1297 more steps per day, when comparing PAM-based interventions with control interventions in 21 studies. Furthermore, this review demonstrates moderate quality of evidence for a small to moderate effect on MVPA time equivalent to 8 more minutes per day. This review did not find an effect on physical capacity, BMI or HRQoL. Given the heterogeneity of the study samples, the results are likely to be applicable to a broad older population, but medication and disease specific treatments need to be considered.

Implications for future practice and research

It seems safe and feasible to use PAMs in PA interventions in older adults. To avoid publication bias and unexplained heterogeneity, more randomized studies with high methodological quality and large sample sizes, are needed to determine possible participant characteristics associated with the adherence to and effect of the interventions. Furthermore, future studies should investigate if PAMs should be included as add-on interventions, or if other types of behavior change strategies should be applied to PAM-based interventions. The evolution of Internet of Things in medicine will emerge and have a great impact on how clinical decision making, preventive medicine and rehabilitation will take place in the future [52, 53]. To ensure that the costs and expenses are used correctly, it seems highly important to have ongoing reviewing of the literature and to include recent published RCTs in updated version of systematic reviews in this area of behavioral intervention research.

Summary box (bullets)

- PAM-based interventions seem to be safe and effective in enhancing the level of PA in older adults.
- Low quality of evidence exists for PAM-based interventions having a moderate effect on PA, equivalent to 1297 more steps per day.
- Moderate quality of evidence exists for PAM-based interventions having a small to moderate effect on MVPA time equivalent to 8 more minutes per day.
- This review could not demonstrate an effect of PAM intervention on physical capacity, BMI or HRQoL.
- Future studies should not use non-active control interventions but instead compare PAM-based interventions with other active interventions or

conduct add-on designs to investigate if the effect size of the PAM-intervention can be increased.

Additional files

Additional file 1: Figure S6. Subgroup analysis on effect of the interventions on physical activity sorted on type of physical activity monitor, diagnoses, feedback frequency, risk of bias judgement and type of physical activity measure. Results are from random effects model using Hedges g. K: Number of studies; SMD: standardized mean difference; PAM: physical activity monitor; COPD: chronic obstructive pulmonary disease. For each analysis, the diamond represents the standardized mean difference of the pooled intervention effect with the horizontal line representing 95% confidence intervals. **Figure S7.** Subgroup analysis on effect of the interventions on moderate to vigorous physical activity, sorted on type of physical activity monitor, diagnoses, feedback frequency, and risk of bias judgement. Results are from random effects model using Hedges g. K: Number of studies; SMD: standardized mean difference; PAM: physical activity monitor; COPD: chronic obstructive pulmonary disease. For each analysis, the diamond represents the standardized mean difference of the pooled intervention effect with the horizontal line representing 95% confidence intervals. **Figure S8.** Subgroup analysis on effect of the interventions on physical capacity, sorted on type of physical activity monitor, diagnoses, feedback frequency, and risk of bias judgement. Results are from random effects model using Hedges g. K: Number of studies; SMD: standardized mean difference; PAM: physical activity monitor; COPD: chronic obstructive pulmonary disease. For each analysis, the diamond represents the standardized mean difference of the pooled intervention effect with the horizontal line representing 95% confidence intervals. **Figure S9.** Subgroup analysis on effect of the interventions on body mass index, sorted on type of physical activity monitor, diagnoses, feedback frequency, and risk of bias judgement. Results are from random effects model using Hedges g. K: Number of studies; SMD: standardized mean difference; PAM: physical activity monitor; COPD: chronic obstructive pulmonary disease. For each analysis, the diamond represents the standardized mean difference of the pooled intervention effect with the horizontal line representing 95% confidence intervals. **Figure S10.** Subgroup analysis on effect of the interventions on health-related qualify of life, sorted on type of physical activity monitor, diagnoses, feedback frequency, and risk of bias judgement. Results are from random effects model using Hedges g. K: Number of studies; SMD: standardized mean difference; PAM: physical activity monitor; COPD: chronic obstructive pulmonary disease; HRQoL: Health-related quality of life. For each analysis, the diamond represents the standardized mean difference of the pooled intervention effect with the horizontal line representing 95% confidence intervals. Positive values favor the intervention. **Figure S11.** Funnel plot with Eggers line illustrating risk of publication bias in the analysis of effect of the interventions on physical activity. SMD: standardized mean difference. **Figure S12.** Random effects meta-analysis on withdrawals due to illness and adverse events. For each study, the diamond represents the specific relative risk of withdrawing with the horizontal line representing 95% confidence intervals. Results are from random effects model with relative risks. RR: Relative risk. The large diamond represents the pooled relative risk. Values below one equals more events in the intervention groups. **Figure S13.** Explorative subgroup analyses of effect of interventions on physical activity sorted on control intervention. For each study, the diamond represents the standardized mean difference of the intervention effect with the horizontal line representing 95% confidence intervals. Results are from random effects model using standardized mean difference (SMD) adjusted to Hedges g. PA: physical activity. The large diamonds represent the pooled standardized mean difference between the intervention groups and the control groups. Positive values favor the intervention. **Figure S14.** Explorative subgroup analyses of effect of interventions on physical activity sorted on additional intervention content. Results are from random effects model using standardized mean difference (SMD) adjusted to Hedges g. For each study, the diamond represents the standardized mean difference of the intervention effect with the horizontal line representing 95% confidence intervals. The large diamonds represent the pooled standardized mean difference between the intervention groups and the control groups. Positive values favor the intervention. **Figure S15.** Figure S15. Explorative subgroup analyses of effect of interventions on physical activity sorted on active control intervention or non-active control intervention. Results are from random effects model using standardized mean difference (SMD) adjusted to Hedges g. For each study, the diamond represents the standardized mean difference of the intervention effect with the horizontal line representing 95% confidence intervals. The large diamonds represent the pooled standardized mean difference between the intervention groups and the control groups. **Table S1.** Characteristics of included studies. **Table S2.** Univariate meta-regressions between standardized mean differences from all outcomes and age, gender distribution, number of participants with walking aids, intervention length, baseline physical activity and body mass index. **Table S3.** Citations and reasons for exclusion from full text screening. (DOCX 4040 kb)

Additional file 2: Summary of findings table. (DOCX 17 kb)

Abbreviations

6MWT: Six-minute walking test; 95%CI: 95% confidence interval; BMI: Body mass index; HRQoL: Health-related quality of life; MVPA: Moderate to vigorous physical activity; PA: Physical Activity; PAMs: Physical activity monitors; RCT: Randomized controlled trial; SD: Standard deviation; SMD: Standard mean difference

Acknowledgements

None.

Funding

The content presented within this paper was produced as part of the project REACH: this project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 690425 [1, 54].

Availability of data and materials

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

Authors' contributions

RTL lead the systematic review and wrote the article manuscript. HBA, HL and RTL participated in the preliminary design of the study. JC, CBJ and RTL designed the search strategy, conducted the inclusion process, risk of bias assessment and data extraction. CBJ and RTL designed the 'Data collection and analysis' section. All authors participated in the writing process. All authors approved the manuscript and agreed to be accountable for all aspects of the work.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]CopenRehab, Department of Public Health, Section of Social Medicine, University of Copenhagen, Gothersgade 160, 3rd floor, 1123 Copenhagen K, Denmark. [2]Department of Occupational- and Physiotherapy, Copenhagen University Hospital, Copenhagen, Denmark. [3]Technical University of Denmark, DTU Management Engineering Institute, Diplomvej 372 office 226, 2800 Lyngby, Denmark. [4]Research Unit of Musculoskeletal Function and Physiotherapy, Institute of Sports Science and Clinical Biomechanics, Faculty

of Health Sciences, University of Southern Denmark, Odense, Denmark. [5]Department of Physiotherapy and Occupational Therapy, Copenhagen University Hospital, Herlev and Gentofte, Denmark. [6]National Centre for Rehabilitation and Palliative Care, University of Southern Denmark and Odense University Hospital, Odense, Denmark.

Received: 19 December 2018 Accepted: 8 April 2019
Published online: 04 May 2019

References

1. Lee I-M, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet Lond Engl. 2012; 380:219–29.
2. Takagi D, Nishida Y, Fujita D. Age-associated changes in the level of physical activity in elderly adults. J Phys Ther Sci. 2015;27:3685–7.
3. McGuire LC, Ahluwalia IB, Strine TW. Chronic disease-related behaviors in U. S. older women: Behavioral Risk Factor Surveillance System, 2003. J Womens Health. 2006;15:3–7.
4. Mummery WK, Kolt G, Schofield G, McLean G. Associations between physical activity and other lifestyle behaviors in older New Zealanders. J Phys Act Health. 2007;4:411–22.
5. Sun F, Norman IJ, While AE. Physical activity in older people: a systematic review. BMC Public Health. 2013;13:449.
6. Gomes M, Figueiredo D, Teixeira L, Poveda V, Paúl C, Santos-Silva A, et al. Physical inactivity among older adults across Europe based on the SHARE database. Age Ageing. 2017;46:71–7.
7. Chodzko-zajko WJ, Proctor DN, Singh MAF, Minson CT, Nigg CR, Salem GJ, et al. Exercise and physical activity for older adults. Med Sci Sports Exerc. 2009;41:1510–30.
8. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA. 2007;298:2296–304.
9. Batsis JA, Naslund JA, Gill LE, Masutani RK, Agarwal N, Bartels SJ. Use of a wearable activity device in rural older obese adults: a pilot study. Gerontol Geriatr Med. 2016;2:233372141667807.
10. McMahon SK, Lewis B, Oakes M, Guan W, Wyman JF, Rothman AJ. Older adults' experiences using a commercially available monitor to self-track their physical activity. JMIR MHealth UHealth. 2016;4:e35.
11. Rowley TW, Lenz EK, Swartz AM, Miller NE, Maeda H, Strath SJ. Efficacy of an individually tailored, internet-mediated physical activity intervention in older adults: a randomized controlled trial. J Appl Gerontol. 2017: 733464817735396.
12. Tabak M, Vollenbroek-Hutten MM, van der Valk PD, van der Palen J, Hermens HJ. A telerehabilitation intervention for patients with chronic obstructive pulmonary disease: a randomized controlled pilot trial. Clin Rehabil. 2014;28:582–91.
13. Mutrie N, Doolin O, Fitzsimons CF, Grant PM, Granat M, Grealy M, et al. Increasing older adults' walking through primary care: results of a pilot randomized controlled trial. Fam Pract. 2012;29:633–42.
14. Peel NM, Paul SK, Cameron ID, Crotty M, Kurrle SE, Gray LC. Promoting activity in geriatric rehabilitation: a randomized controlled trial of accelerometry. PLoS One. 2016;11:e0160906.
15. Nishiguchi S, Yamada M, Tanigawa T, Sekiyama K, Kawagoe T, Suzuki M, et al. 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. 2015;63:1355–63.
16. Furber S, Butler L, Phongsavan P, Mark A, Bauman A. Randomised controlled trial of a pedometer-based telephone intervention to increase physical activity among cardiac patients not attending cardiac rehabilitation. Patient Educ Couns. 2010;80:212–8.
17. Higgins J, Green S. Cochrane handbook for systematic reviews of interventions. 2011. http://handbook.cochrane.org/. Accessed 11 Apr 2016.
18. Larsen RT, Christensen J, Juhl C, Andersen HB, Langberg H. PROSPERO register: physical activity monitors to enhance the daily amount of steps taken in the elderly – a protocol for a systematic review and meta-analysis. PROSPERO. 2018. http://www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42018083648. Accessed 26 Apr 2018.
19. Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance the daily amount of physical activity in elderly—a protocol for a systematic review and meta-analysis. Syst Rev. 2018;7:69.
20. Higgins J, Green S. 16.4.5 Methods for incorporating cross-over trials into a meta-analysis. Version 5.1.0. Cochrane Collaboration; 2011. http://handbook.cochrane.org/chapter_16/16_4_6_1_mean_differences.htm. Accessed 3 Oct 2016.
21. Higgins J, Savovic J, Sterne JAC, et al. A revised tool to assess risk of bias in randomized trials (RoB 2.0). 2016.
22. Bliddal H, Christensen R. The treatment and prevention of knee osteoarthritis: a tool for clinical decision-making. Expert Opin Pharmacother. 2009;10:1793–804.
23. Yates T, Davies M, Gorely T, Bull F, Khunti K. Effectiveness of a pragmatic education program designed to promote walking activity in individuals with impaired glucose tolerance: a randomized controlled trial. Diabetes Care. 2009;32:1404–10.
24. Pelssers J, Delecluse C, Opdenacker J, Kennis E, Van Roie E, Boen F. "Every step counts!": effects of a structured walking intervention in a community-based senior organization. J Aging Phys Act. 2013;21:167–85.
25. Kolt GS, Schofield GM, Kerse N, Garrett N, Ashton T, Patel A. Healthy steps trial: pedometer-based advice and physical activity for low-active older adults. Ann Fam Med. 2012;10:206–12.
26. Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. Stat Med. 2002;21:1575–600.
27. Higgins JPT, Green S. 'Cochrane Handbook for Systematic Reviews of Interventions, 9.4.6 Combining Dichotomous and Continuous Outcomes'. In Cochrane Handbook for Systematic Reviews of Interventions, 2011. http://handbook-5-1.cochrane.org/chapter_9/9_4_6_combining_dichotomous_and_continuous_outcomes.htm.
28. Higgins G. Cochrane Handbook for Systematic Reviews of Interventions, 16. 5.4 How to include multiple groups from one study. In: Cochrane Handbook for Systematic Reviews of Interventions; 2011.
29. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000; 56:455–63.
30. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315:629–34.
31. Nolan CM, Maddocks M, Canavan JL, Jones SE, Delogu V, Kaliaraju D, et al. Pedometer step count targets during pulmonary rehabilitation in chronic obstructive pulmonary disease. a randomized controlled trial. Am J Respir Crit Care Med. 2017;195:1344–52.
32. McMurdo MET, Sugden J, Argo I, Boyle P, Johnston DW, Sniehotta FF, et al. Do pedometers increase physical activity in sedentary older women? A randomized controlled trial. J Am Geriatr Soc. 2010;58:2099–106.
33. Croteau KA, Richeson NE, Farmer BC, Jones DB. Effect of a pedometer-based intervention on daily step counts of community-dwelling older adults. Res Q Exerc Sport. 2007;78:401–6.
34. Croteau KA, Richeson NE, Vines SW, Jones DB. Effects of a pedometer-based physical activity program on older adults' mobility-related self-efficacy and physical performance. Act Adapt Aging. 2004;28:19–33.
35. Harris T, Kerry SM, Victor CR, Ekelund U, Woodcock A, Iliffe S, et al. A primary care nurse-delivered walking intervention in older adults: PACE (pedometer accelerometer consultation evaluation)-lift cluster randomised controlled trial. PLoS Med. 2015;12:e1001783.
36. Kawagoshi A, Kiyokawa N, Sugawara K, Takahashi H, Sakata S, Satake M, et al. Effects of low-intensity exercise and home-based pulmonary rehabilitation with pedometer feedback on physical activity in elderly patients with chronic obstructive pulmonary disease. Respir Med. 2015;109:364–71.
37. Merom D, Grunseit A, Eramudugolla R, Jefferis B, McNeill J, Anstey KJ. Cognitive benefits of social dancing and walking in old age: The dancing mind randomized controlled trial. Front Aging Neurosci. 2016;8:26.
38. Sugden JA, Sniehotta FF, Donnan PT, Boyle P, Johnston DW, McMurdo MET. The feasibility of using pedometers and brief advice to increase activity in sedentary older women - A pilot study. BMC Health Serv Res. 2008;8:169.
39. Talbot LA, Gaines JM, Huynh TN, Metter EJ. A home-based pedometer-driven walking program to increase physical activity in older adults with osteoarthritis of the knee: a preliminary study. J Am Geriatr Soc. 2003;51:387–92.
40. Lee L-L, Arthur A, Avis M. Evaluating a community-based walking intervention for hypertensive older people in Taiwan: a randomized controlled trial. Prev Med. 2007;44:160–6.
41. Koizumi D, Rogers NL, Rogers ME, Islam MM, Kusunoki M, Takeshima N. Efficacy of an accelerometer-guided physical activity intervention in community-dwelling older women. J Phys Act Health. 2009;6:467–74.

42. McLellan AG, Slaght J, Craig CM, Mayo A, Sénéchal M, Bouchard DR. Can older adults improve the identification of moderate intensity using walking cadence? Aging Clin Exp Res. 2018;30:89–92.
43. World Health Organization. Global recommendations on physical activity for health. 2010. http://www.ncbi.nlm.nih.gov/books/NBK305057/. Accessed 26 Jun 2017.
44. Tudor-Locke C, Craig CL, Aoyagi Y, Bell RC, Croteau KA, De Bourdeaudhuij I, et al. How many steps/day are enough? For older adults and special populations. Int J Behav Nutr Phys Act. 2011;8:80.
45. Qiu S, Cai X, Ju C, Sun Z, Yin H, Zügel M, et al. Step counter use and sedentary time in adults. Medicine (Baltimore). 2015;94. https://doi.org/10.1097/MD.0000000000001412.
46. Koolhaas CM, van Rooij FJA, Schoufour JD, Cepeda M, Tiemeier H, Brage S, et al. Objective measures of activity in the elderly: distribution and associations with demographic and health factors. J Am Med Dir Assoc. 2017;18:838–47.
47. Weiss CO. Frailty and chronic diseases in older adults. Clin Geriatr Med. 2011;27:39–52.
48. Solé-Auró A, Michaud P-C, Hurd M, Crimmins E. Disease incidence and mortality among older Americans and Europeans. Demography. 2015;52:593–611.
49. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions, 10.2 Types of reporting bias and the supporting evidence. In: Cochrane Handbook for Systematic Reviews of Interventions; 2011.
50. Karlsson P, Bergmark A. 'Compared with What? An Analysis of Control-Group Types in Cochrane and Campbell Reviews of Psychosocial Treatment Efficacy with Substance Use Disorders: Compared with What?' Addiction 110, no. 3. 2015:420–28. https://doi.org/10.1111/add.12799.
51. Waters L, Reeves M, Fjeldsoe B, Eakin E. Control group improvements in physical activity intervention trials and possible explanatory factors: a systematic review. J Phys Act Health. 2012;9:884–95.
52. Dimitrov DV. Medical internet of things and big data in healthcare. Healthc Inform Res. 2016;22:156–63.
53. Wright SP, Hall Brown TS, Collier SR, Sandberg K. How consumer physical activity monitors could transform human physiology research. Am J Physiol - Regul Integr Comp Physiol. 2017;312:R358–67.
54. REACH | HORIZON 2020. http://reach2020.eu/?page_id=1485. Accessed 23 Nov 2017.

APPENDIX 3: PAPER C (STUDY II)

European Review of Aging and Physical Activity

RESEARCH ARTICLE Open Access

Criterion validity for step counting in four consumer-grade physical activity monitors among older adults with and without rollators

Rasmus Tolstrup Larsen[1*], Christoffer Brun Korfitsen[1], Carsten Bogh Juhl[2,3], Henning Boje Andersen[4], Henning Langberg[1] and Jan Christensen[5,6]

Abstract

Background: Few studies have investigated the measurement properties of consumer-grade physical activity monitors (PAMs) in older adults. Therefore, we investigated the criterion validity of consumer-grade PAMs in older adults and whether the measurement properties differed between older adults with and without rollators and whether worn on the hip or at the wrist.

Methods: Consumer-grade PAMs were eligible for inclusion in this study if they: 1) could be fastened at the hip as well as on the wrist, 2) were simple in function and design and thus easy to use for participants with minimal technical skills, 3) included step-counting as outcome measure and 4) were powered by a button cell battery. Participants performed self-paced walking for six minutes while two physiotherapists counted their steps with a click-counter. The average of the two counts was used as criterion. The participants wore 16 monitors, four located bilaterally on both hips and wrists. Our prior expectation was that all monitors would have at least moderate criterion validity for all participants, good criterion validity for participants walking without a rollator and poor criterion validity for participants walking with a rollator.

Results: Four physical activity monitors were included in this study; Misfit Shine, Nokia GO, Jawbone UP Move and Garmin Vivofit 3. A total of 103 older adults participated.
Nokia GO was excluded from this study due to technical issues. Therefore, we present results on the frequency of data loss, ICC (1, 2) and percentage measurement error for Misfit Shine, Garmin Vivofit 3 and Jawbone UP Move located on four different positions.

Conclusions: The hip-worn PAMs did not differ significantly in terms of measurement error or criterion validity. Wrist-worn monitors cannot adequately measure number of steps in a population of older adults using rollators. The hip-worn PAMs were superior to wrist-worn PAMs among older adults with and without rollators.

Keywords: Validity, Physical activity monitors, Walking, Technology

* Correspondence: rala@sund.ku.dk
[1]Department of Public Health, Faculty of Health and Medical Sciences, CopenRehab, Section of Social Medicine, University of Copenhagen, Gothersgade 160, 3rd floor, 1123 Copenhagen K, Denmark
Full list of author information is available at the end of the article

Background

Functional decline is related to aging. Still, older adults who engage in exercise or physical activity regularly can, to some extent, maintain their physical function, have lower all-cause mortality, are less disabled, and have a lower prevalence of several non-communicable diseases [1–4]. Walking is the favourite activity among community-dwelling cognitively-intact older adults [5]. Furthermore, walking programmes have in several systematic reviews been shown to be effective in increasing physical activity in the short term in older adults [6]. However, to ensure long-lasting effects and adherence of walking programs, they should be individualised and based on behavioural theories, as well as include goals to maintain acceptable levels of PA [6].

To use goal setting in the individualisation of walking programs, individual feedback on PA is crucial. The consumer-grade physical activity monitors (PAMs) hold the potential of being a facilitator for increased PA as they provide timed feedback, notifications and can be adjusted with individual goals [7]. For these reasons, PAMs are now frequently used with good effect to increase physical activity in older adults [8, 9]. However, before using consumer-grade PAMs in clinical research, the measurement properties, including criterion validity in particular, of specific PAMs should be evaluated [10]. Measurement properties for specific PAMs may differ between different populations of older adults. Thus, it has been shown that adults suffering from knee pain or those who depend on a walker have different gait characteristics compared with normal older adults [11]. Within the population of older adults, a large heterogeneity exists in gait speed, stride length, joint movement, and use of assistive devices, all of which have been found to affect the validity of PAMs [11, 12].

Furthermore, consumer-grade PAMs differ from research-grade PAMs because the algorithms for step detection cannot be modified and thus the definition of a step might differ between PAMs. Hence, there is no transparency in the use of algorithms. Besides, most modern consumer-grade PAMs are designed to be worn on the wrist as watches, which might lead to inaccurate measurement as hip-worn PAMs have been reported to outperform wrist-worn PAMs for step accuracy [13].

To our knowledge, few studies have investigated the measurement properties of consumer-grade PAMs in older adults, and none of these has studied the measurement properties of a given PAM model worn on the hip and wrist [12, 14–20]. Therefore, the present study aimed to investigate (a) the criterion validity of four consumer-grade PAMs in older adults performing a self-paced indoor walking test and (b) whether the measurement properties of the PAMs differed between older adults with and without rollators and comparing wrist-worn and hip-worn positions.

Methods

Participants

We included older adults from five community activity centres in the municipality of Copenhagen, Denmark. The participants were recruited at the 'morning meet-up' where our research team presented the study. Participants were eligible if they were 65 years or above, community-dwelling, living at home and able to walk independently with or without a rollator or cane. Mild and more severe cognitive impairment was an exclusion criterion, since participants had to be able to understand the study aims and fill out themselves the baseline questions.

Ethics

Oral and written information was given before participants gave informed consent to participate. The study was approved by the Danish Ethics Committee (Journal nr.:H-17033310).

Physical activity monitors

As we could not investigate all available PAMs, we chose those who were most relevant for older adults and those who allowed us to investigate whether the placement of the specific PAM affected the validity. Thus, consumer-grade PAMs were eligible for inclusion in this study if they: 1) could be fastened at the hip as well as on the wrist, 2) were simple in function and design and requiring no technical skills to be operated, 3) included step-counting as the outcome measure and 4) powered by a button cell battery providing a battery life for more than three months. If the included PAMs did not have a display, they were paired with an iPod Touch 5th generation, model A1421, operating with iOS 9.3.5. We performed pilot testing of all the eligible consumer-grade PAMs within the research team before conducting the present study.

Procedures and measures

Participants were included between March and June 2018. In the five activity centres, participants were asked to perform self-paced walking for six minutes. To secure the external validity of our results, we asked the participants were asked to walk at their normal gait speed, instead of a maximal walking test.

An unobstructed 15- or 30-m flat track was used for testing, at each end a cone was positioned indicating where participants should make a 180-degree turn. The participants decided themselves whether they performed right or left turns. If the participants were interrupted during the testing or became tired, they were allowed to rest standing or sitting and the time was stopped until they continued. A chair was provided upon request. The participants received no verbal feedback on gait speed from the testers. The participants were fitted with 16 PAMs (four models located bilaterally on both hips and wrists). The hip-worn monitors were fitted to the belt of the participant or to front pocket sewing, the wrist-worn monitors were fitted with the rubber straps provided, and in both cases testers assisted with fitting.

The order of the PAMs was changed between every participant to ensure a balanced order throughout the study. Anthropometric measures of weight and height and demographic data and information of health-related behaviors were obtained prior to the test session. During each test walks, two physiotherapists were positioned by each cone and, blinded from the other physiotherapist's counting, counted the steps with a click-counter. The testers were the same for all participants.

Statistical analysis

Normal distributions of continuous data (steps, age, height, body mass index, meters walked in 6 min, and self-paced speed) were evaluated by quantile-quantile plots and histograms of the standardised residuals. Normally distributed continuous data were summarised by means and 95% confidence intervals. Continuous data without a normal distribution were summarised by medians and interquartile ranges. Categorical data were summarised with frequencies and percentage of the total score. The average of the visually counted steps from tester A and tester B was defined as the actual steps taken and hence the criterion. For every participant, four measures for each type of PAM were taken (left hip, right hip, left wrist, and right wrist). The frequency of excluded data points was reported and evaluated between groups with a Chi [2] test.

Interclass correlation coefficients (ICC) were calculated based on a two-way random effects analysis of variance model examining the absolute agreement of a single measure (ICC2,1) [21, 22]. We chose ICC2,1 as the raters were the same, and each participant was rated only once (average between the two testers). The model was chosen to examine the agreement between observed steps and the steps counted by the PAM. ICC (2,1) values of < 0.5, $=0.5- < 0.75$, $=0.75- < 0.9$, and ≥ 0.90 were interpreted as the PAM having, respectively, poor, moderate, good, and excellent criterion validity [21, 23]. Interclass correlation coefficients of mean difference in steps between observed steps and measured steps as well as percent measurement error were reported for 1) all participants, 2) participants without a rollator and 3) participants with a rollator. Our prior expectation was that each of the PAMs, would have at least moderate criterion validity for all participants (but with a low precision of the estimate because of the heterogeneity of the population), a good criterion validity for participants walking without a rollator and a poor criterion validity for participants walking with a rollator (as a previous study has shown that some PAMs have lower measurement properties among rollator users [12]). We expected a better criterion validity in participants without rollators because they were expected to walk faster and more similar to younger adults, compared with participants with rollators.

Visualisation of the absolute percentage measurement errors for each PAM was presented with a scatter plot and analysed with a generalised linear logit link model.

StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC, was used for all statistical analyses and visualisations. An alpha level on 0.05 was considered the threshold for statistical significance.

Results

We identified four consumer-grade wearable PAMs available on the commercial market that met our eligibility criteria: Misfit Shine, Nokia GO, Jawbone UP Move and Garmin Vivofit 3. Below, Fig. 1 a and b show the four included monitors as they were used in this study. Only the Garmin Vivofit 3 included a regular display. The other monitors used light or illustrations to show how close the user is to the step goal of the day. Thus, the Garmin Vivofit 3 was the only monitor that could be operated without a smartphone device for this study.

A total of 103 older adults volunteered to participate in this study. Anthropometric, demographic data and information on health-related behaviour are presented in Table 1.

Deleted observations due to missing data and technical issues

The frequencies of excluded data points due to technical issues are listed in Table 2. We were unable to perform the necessary synchronization of the Nokia GO between each participant; thus, it was not possible to extract data for individual participants from the devices as the Nokia GO does not provide on the PAM itself the number of steps taken. Hence, the Nokia GO devices were excluded from the study. After April 1, 2018, an update to the Misfit iOS application, resulted in a malfunction in the synchronization between the iPod Touch and the Misfit monitors. As a result of this we had to excluded two of the Misfit monitors from that date. The remaining two monitors were positioned on the dominant side of the participants. In total, 103 data points were available for the Garmin and Jawbone monitors, 37 for the left-worn Misfit monitors and 99 for the right-worn Misfit monitors.

Fig. 1 and 2 illustrates the percentage of excluded data points. In total, there were 175 excluded data points

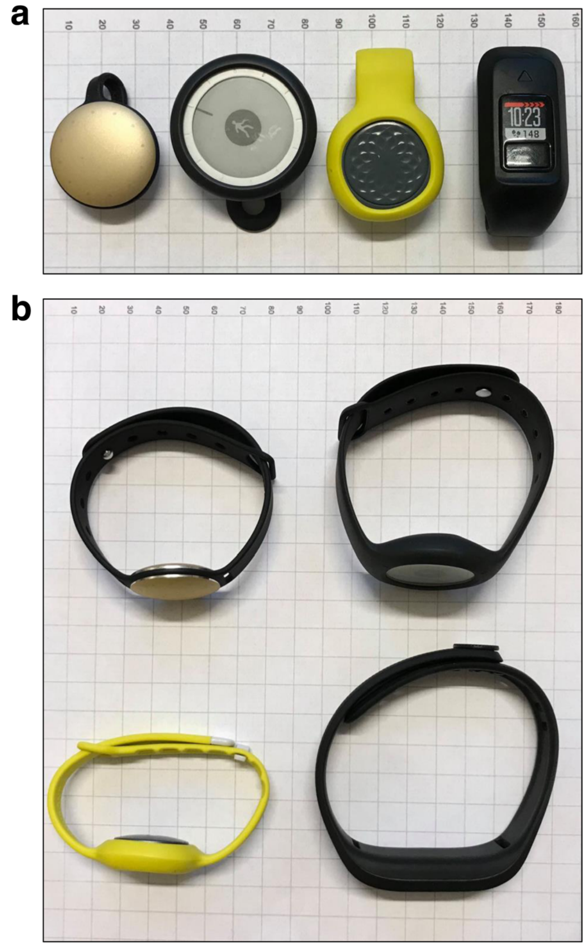


Fig. 1 **a** and **b** From left to right: Misfit Shine, Nokia GO, Jawbone UP Move and Garmin Vivofit 3 on paper with 10-mm grid lines. Figure a shows the hip-worn physical activity monitors and below figure b shows the wrist-worn physical activity monitors

(16.0%), corresponding to 48 excluded hip measures (27.4%) and 127 excluded wrist measures (72.57%). A Chi [2] test revealed that wrist measures were more likely to be excluded ($p < 0.001$). In total, 8.0% of the Garmin Vivofit 3 measures, 28.2% of the Jawbone UP Move measures, and 9.6% of the Misfit Shine measures were excluded. A Chi [2] test revealed a significant between-group difference (p < 0.001). In total, 16.3% of the left-side measures and 15.7% of the right-side measures were excluded. A Chi [2] test revealed a no between-group difference ($p = 0.816$).

Table 2, reports results on criterion validity ICC (2,1), mean difference and percentage measurement error for all PAMs on all positions. For the hip-worn monitors, 10 out of 18 possible combinations (brand, left/right, and with or without rollator) fulfilled the a priori hypothesis of criterion validity. For the wrist-worn monitors, only one combination fulfilled the a priori hypothesis of criterion validity. The hip-worn Misfit Shine fulfilled four out of six possible combinations of criterion validity (left/right for all participants, participants with rollators and participants without rollators). The hip-worn Garmin Vivofit 3 fulfilled five out of six combinations for criterion validity. The hip-worn Jawbone UP Move fulfilled one out of six combinations for criterion validity. For the wrist-worn PAMs, no combination fulfilled the a priori hypothesis for criterion validity except the right-worn Garmin Vivofit 3 for participants with rollators. Good interrater reliability, ICC (2,1) was found between the two testers 0.88 (95% CI 0.83 to 0.92), with a mean difference on 4.42 steps 95% CI (– 6.10 to – 14.91), (103 measures).

Measurement error

Fig. 3 a, b and c illustrates the relationship between measurement error in percent and observed steps. The

Table 1 Participants characteristics ($n = 103$)

Sex, male, n (%)	35 (34.0%)
Age, mean (95%CI)	81.3 years (79.8 to 82.8)
Height, mean (95%CI)	164.0 cm (162.2 to 165.9)
Body mass index, mean (95%CI)	28.0 kg/m [2] (27.0 to 29.0)
Self-paced meters walked in 6 min, mean (95%CI)	255.0 m (238.5 to 271.4)
Self-paced speed over the 6 min, mean (95%CI)	2.6 km/t (2.4 to 2.7)
Walking without aid, n (%)	52 (50.5%)
Walking with a cane, n (%)	15 (14.5%)
Walking with a rollator, n (%)	36 (35.0%)
Never smoked, n (%)	44 (42.7%)
Stopped smoking, n (%)	48 (46.6%)
Current smoker, n (%)	11 (10.7%)

Abbreviations: 95%CI: 95% Confidence interval; *IQR* Interquartile Range
Normal distributed continuous data: Age, Height, Body Mass Index, Meters walked in 6 min, Self-paced speed over the 6 min

Table 2 A priori hypothesis for criterion validity, criterion validity, mean difference between measured steps and observed steps, and mean percentage measurement error for each physical activity monitor separately for each position

Position and type of monitor	A priori hypothesis ICC(2,1)	ICC(2,1) (95%CI)	Mean difference (95%CI)	Mean % measurement error (95%CI)
Hip-worn Misfit Shine, left				
All (34 measures)	0.5 to 0.75	**0.52 (0.21 to 0.73)**	−61.99 (− 104.50 to − 19.50)	−12.46% (− 21.09 to − 3.83)
Rollator (11 measures)	< 0.5	**0.56 (0.04 to 0.86)**	−73.00 (− 163.94 to 17.94)	−15.87% (− 35.30 to 3.55)
Without rollator (23 measures)	≥ 0.75	0.49 (0.13 to 0.75)	−56.71 (− 107.56 to − 5.88)	− 10.83% (− 20.76 to − 0.90)
Hip-worn Misfit Shine, right				
All (88 measures)	0.5 to 0.75	**0.64 (0.47 to 0.75)**	−48.35 (− 74.47 to − 22.24)	−8.75% (− 14.10 to − 3.40)
Rollator (31 measures)	< 0.5	0.44 (0.08 to 0.69)	− 110.19 (− 169.15 to − 51.24)	− 20.44% (− 31.87 to − 9.02)
Without rollator (57 measures)	≥ 0.75	**0.78 (0.66 to 0.87)**	−14.72 (− 36.11 to 6.65)	− 2.38% (− 7.39 to 2.61)
Hip-worn Garmin Vivofit 3, left				
All (100 measures)	0.5 to 0.75	**0.67 (0.53 to 0.78)**	−41.49 (− 64.21 to − 18.76)	− 9.74% (− 15.16 to − 4.33)
Rollator (36 measures)	< 0.5	**0.57 (0.19 to 0.78)**	−87.5 (− 131.23 to − 43.83)	−20.61% (− 31.02 to − 10.18)
Without rollator (64 measures)	≥ 0.75	0.71 (0.56 to 0.81)	−15.59 (− 39.88 to 8.71)	−3.63% (− 9.42 to 2.16)
Hip-worn Garmin Vivofit 3, right				
All (102 measures)	0.5 to 0.75	**0.80 (0,72 to 0,87)**	−22,61 (− 37.50 to − 7.72)	− 5.18% (− 9.01 to − 1.36)
Rollator (35 measures)	< 0.5	**0.74 (0.45 to 0.87)**	−44.11 (− 70.02 to − 18.21)	− 10.12% (− 16.60 to − 3.64)
Without rollator (67 measures)	≥ 0.75	**0.83 (0.73 to 0.89)**	−11.38 (− 29.46 to 6.70)	−2.61% (− 7.35 to 2.13)
Hip-worn Jawbone UP Move, left				
All (84 measures)	0.5 to 0.75	**0.61 (0.34 to 0.76)**	−63.75 (− 87.94 to − 39.56)	−13.11% (− 18.24 to − 7.98)
Rollator (23 measures)	< 0.5	0.40 (0.00 to 0.72)	− 101.65 (− 144.66 to − 58.64)	−19.21% (− 27.48 to − 10.94)
Without rollator (61 measures)	≥ 0.75	0.64 (0.44 to 0.78)	−49.45 (− 78.44 to − 20.48)	− 10.81% (− 17.19 to − 4.43)
Hip-worn Jawbone UP Move, right				
All (92 measures)	0.5 to 0.75	0.47 (0.21 to 0.65)	−85.79 (− 116.65 to − 54.95)	−16.57% (− 23.02 to − 10.12)
Rollator (31 measures)	< 0.5	0.24 (0.00 to 0.54)	− 193.83 (− 258.89 to − 128.78)	−38.28% (− 51.73 to − 24.84)
Without rollator (61 measures)	≥ 0.75	0.68 (0.51 to 0.80)	− 30.89 (− 54.96 to − 6.83)	−5.53% (− 10.86 to − 0.20)
Wrist-worn Misfit Shine, left				
All (36 measures)	0.5 to 0.75	0.18 (0.00 to 0.46)	−238.43 (−313.06 to − 163.81)	−44.21% (− 57.66 to − 30.78)
Rollator (12 measures)	< 0.5	0.00 (0.00 to 0.07)	− 486.5 (− 568.45 to − 404.55)	−91.03% (− 95.79 to − 86.27)
Without rollator (24 measures)	≥ 0.75	0.37 (0.00 to 0.68)	− 114.40 (− 170.90 to − 57.89)	− 20.80% (− 31.65 to − 9.96)
Wrist-worn Misfit Shine, right				
All (88 measures)	0.5 to 0.75	0.23 (0.00 to 0.47)	− 220.38 (− 266.13 to −174.64)	− 41.91% (− 50.49 to − 33.34)
Rollator (30 measures)	< 0.5	0.02 (0.00 to 0.09)	− 462.83 (− 518.31 to − 407.34)	−89.03% (− 97.04 to − 81.02)
Without rollator (58 measures)	≥ 0.75	0.55 (0.10 to 0.77)	− 94.97 (− 124.91 to − 65.03)	− 17.55% (− 23.39 to − 11.71)
Wrist-worn Garmin Vivofit 3, left				
All (88 measures)	0.5 to 0.75	0.31 (0.06 to 0.52)	− 139.71 (−186.39 to − 93.05)	− 27.17% (− 36.14 to − 18.20)
Rollator (22 measures)	< 0.5	0.00 (0.00 to 0.08)	− 455.78 (− 524.28 to − 387.27)	−88.31% (− 98.97 to − 77.67)
Without rollator (66 measures)	≥ 0.75	0.67 (0.51 to 0.79)	−34.36 (− 61.45 to − 7.28)	−6.79% (− 12.43 to −1.15)
Wrist-worn Garmin Vivofit 3, right				
All (89 measures)	0.5 to 0.75	0.33 (0.08 to 0.53)	− 132.98 (− 179.05 to − 86.91)	− 26.47% (− 35.62 to − 17.33)
Rollator (23 measures)	< 0.5	0.01 (0.00 to 0.08)	− 455.17 (− 519.00 to − 391.34)	−88.98% (− 98.99 to − 78.96)
Without rollator (66 measures)	≥ 0.75	**0.76 (0.63 to 0.84)**	−20.70 (− 42.4786 to 1.069512)	−4.69% (− 10.21 to 0.82)
Wrist-worn Jawbone UP Move, left				
All (65 measures)	0.5 to 0.75	0.30 (0.03 to 0.52)	− 121.52 (−166.86 to − 76.19)	−21.87% (− 30.14 to − 13.61)
Rollator (7 measures)	< 0.5	0.01 (0.00 to 0.23)	− 480.86 (− 640.08 to − 321.64)	−84.97% (− 110.58 to − 59.38)

Table 2 A priori hypothesis for criterion validity, criterion validity, mean difference between measured steps and observed steps, and mean percentage measurement error for each physical activity monitor separately for each position *(Continued)*

Position and type of monitor	A priori hypothesis ICC(2,1)	ICC(2,1) (95%CI)	Mean difference (95%CI)	Mean % measurement error (95%CI)
Without rollator (58 measures)	≥ 0.75	0.47 (0.18 to 0.68)	−78.15 (− 112.13 to − 44.18)	−14.25% (− 20.81 to − 7.70)
Wrist-worn Jawbone UP Move, right				
All (55 measures)	0.5 to 0.75	0.29 (0.02 to 0.53)	− 105.05 (−148.00 to − 62.11)	−18.89% (− 26.65 to − 11.13)
Rollator (3 measures)	< 0.5	0.00 (0.00 to 0.88)	− 386.0 (− 1157.00 to 384.91)	− 66.17% (− 195.37 to 63.01)
Without rollator (52 measures)	≥ 0.75	0.38 (0.06 to 0.61)	−88.84 (− 126.23 to − 51.46)	−16.16% (− 23.13 to − 9.19)

Criterion validity calculated using a two-way random, single measures, absolute agreement model end expressed as interclass correlation coefficient
Abbreviations; *ICC* Interclass Correlation Coefficient (bold equal fulfilling the a priori hypothesis), *MD* Mean Difference: 95% Confidence intervals
Measurement error in % were evaluated as being not normally distributed and are presented with median and interquartile range. ICC (2, 1) values that meet the a priori hypothesis are marked with bold

logit link models reveal a negative slope for all PAMs in participants without rollators and for hip-worn monitors for participants with rollators. The models for wrist-worn monitors in participants with rollators differ from for the other models as the slope is more horizontal and has larger measurement error. There is no visual difference between any left and right measures.

Discussion

The aim of this study was to investigate the criterion validity of four types of consumer-grade PAMs in older adults. The loss of data due to technical issues is more likely to happen with wrist-worn monitors. The Garmin Vivofit 3 showed the lowest frequency of lost datapoints and the Nokia GO was excluded from the study being incapable of synchronizing data. This means that we cannot rule out the Nokia GO as a PAM with acceptable measurement properties, as it might work very well with other devices. However, to be as transparent as possible, we chose to describe the Nokia GO with the same detail as the other PAMs. Hip-worn PAMs were superior to wrist-worn PAMs across all participants, participants without and with rollators in terms of criterion validity, absolute difference in steps, absolute measurement error in percentage and difference in steps.

Loss of data due to technical issues is often reported among consumer-grade PAMs [20]. In this study, none of the investigated PAMs was free from data loss but some of the PAMs were clearly more affected by this problem than others. Excluding lost data and zero

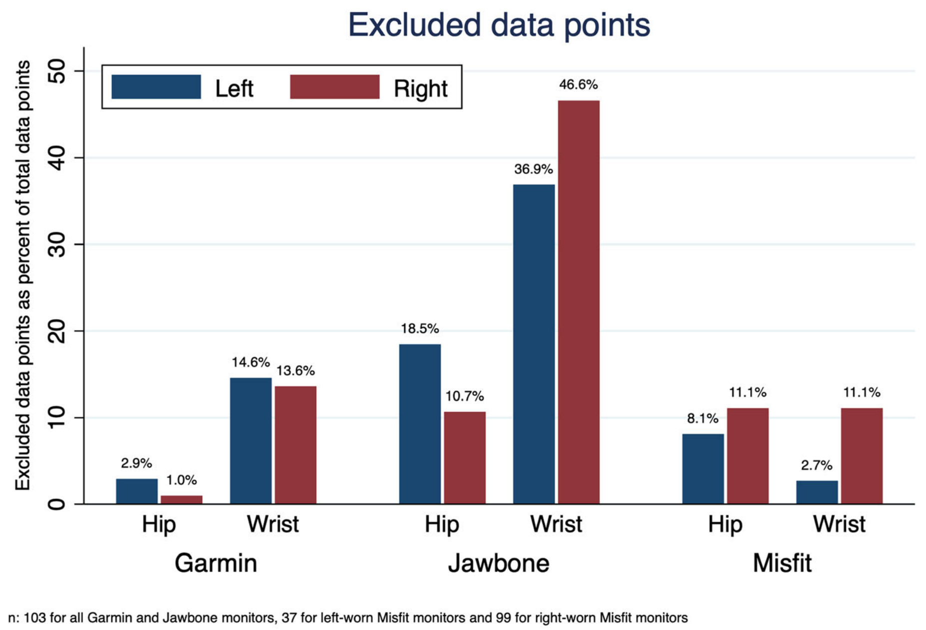


Fig. 2 Excluded data points as a percentage of total data points sorted on brand and position. Higher percentage equals more excluded data. Chi [2]-tests revealed a significant difference between brands ($p < 0.001$) hip and wrist positions ($p < 0.001$) but not between left and right positions ($p = 0.816$)

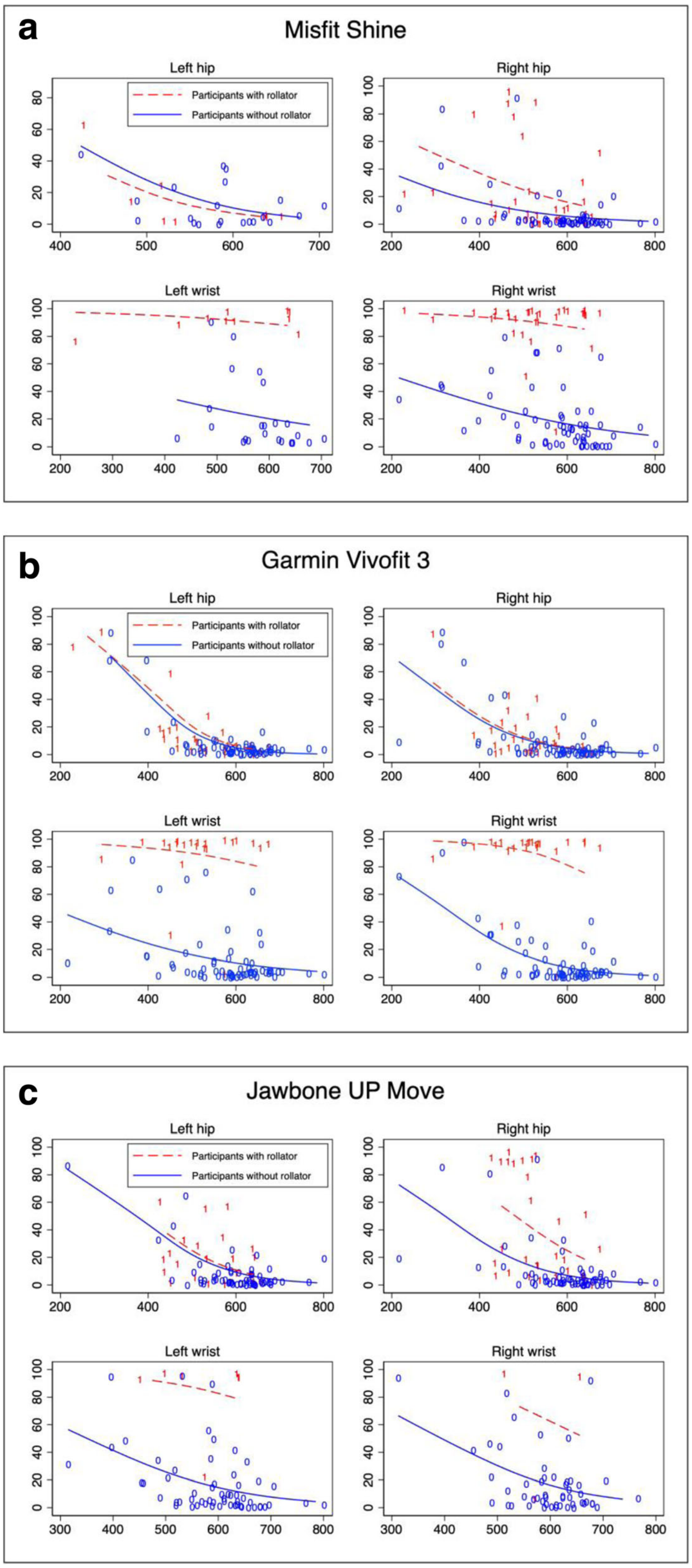


Fig. 3 (See legend on next page.)

(See figure on previous page.)
Fig. 3 a, **b** and **c** Two-way scatter plots with logit link generalized linear models between absolute measurement error in % and observed steps (criterion) of Misfit Shine (a), Garmin Vivofit 3 (b) and Jawbone UP Move (c) physical activity monitors. Each figure includes results from the left hip, right hip, left wrist and right wrist. Red digits "1" and lines equal participants with rollators and blue digits "0" and lines equal participants without rollators. Y-axis represent absolute measurement error in % as a response to the x-axis which is number of observed steps

counts will affect the criterion validity and cause a systematically higher interclass correlation compared to analysis with included zero counts. The interpretation of the ICC (2,1) value cannot stand alone and when evaluating the measurement properties of a PAM, results on data loss should be interpreted as well. Fig. 2 illustrates the problem in each brand, position and body side. The Garmin Vivofit 3 monitor and the Misfit Shine monitor had the lowest affection of data loss, but we had to exclude two of the Misfit monitors halfway, reducing the precision of our results. It also illustrates that wrist measures were more likely to be excluded, as many of the measures did not count when participants were using a rollator, and similarly, that there was no difference in exclusion of data between left and right-side measures.

The logit link models from Fig. 3 illustrate the relationship between measurement error and observed steps among participants with and without rollators. For the hip-worn PAMs among all participants and for wrist-worn PAMs among participants without rollators, the relationship was similar. In line with several other studies of consumer-grade PAMs in older adults, we found a higher accuracy in faster walking older adults [17, 19, 20, 24]. As described in the introduction, walkers with assistive devices are more likely to have alternative gait pattern compared to walkers using no assistive device. For participants using a rollator, the horizontal logit link models showed close to 100% absolute measurement error in wrist-worn PAMs indicating lack of arm movement among rollator users.

In terms of statistical methods, we chose to analyse the primary outcome using the two-way random effects model with absolute agreement and single measures, ICC (2, 1). In this model, each tester measures each participant, and testers are considered representative of a larger population of testers. Previously studies have either used Pearson correlation coefficients [19], unspecified ICC [17, 18] or ICC (2, 1) [12, 14, 15]. Agreement between two continuous outcomes should be reported using ICC values [25], and future studies should as a minimum report the specific sub-type of ICC as well as difference (percentage or mean) allowing the results to be compared between studies.

The criterion represents the actual true number of steps taken. When visually counting the steps, we avoided technical solutions of counting steps for the criterion. Other papers have often used research-grade accelerometers to validate consumer-grade PAMs [12, 14–19] which is the best option for free-living conditions. However, strictly for walking, the validity of research-grade PAMs can be questioned in this population as consumer-grade PAMs have been reported to have greater validity in trials comparing them to research-grade PAMs against visually counted steps [15]. With complex gait patterns in populations containing participants with and without walking aids the visually counted number of steps must serve as the most valid criterion, which was why we chose this method and in contrast to other studies with visual counts, we tried to reduce counting bias by having two testers instead of only one [15]. To exclude all error from the criterion, we could have combined more testers but it was not possible in this setting. However, all methods will have flaws and since there was no significant difference between the counts of the testers, we should be able to trust the average as a true criterion.

This study holds several limitations in the interpretation of the results. Firstly, the results are only generalisable to self-paced indoor walking in older adults. A study by Grant et al. reported large differences between counts from some research-grade PAMs in indoor treadmill walking and outdoor walking, but only in the slowest walking speeds [26]. To our knowledge, no published similar comparison has been made in free walking and using consumer-grade PAMs, but this highlights the lack of evidence in this area. Furthermore, the approach used for this study was general and covers only cyclic gait. The outcome of interest was step count when walking and did not include specific movements such as turning or squatting. Thus, our results only cover validity in cyclic gait and these results cannot be generalised and should not be extrapolated to conclude upon accelerometery vector counts in more specific movements. To investigate this, the raw data from the consumer-grade PAMs must be available for researchers, and until then, consumer-grade PAMs still remain as "black boxes" with hidden filtering software.

Secondly, we cannot rule out the possibility of existing PAMs, fulfilling our inclusion criteria that we were not aware of. We searched the literature and the web pages of all the major brands for relevant PAMs, but in the end our results do not apply other PAMs than the four devices we included in this study.

Another limitation is the possible systematic error in our dataset due to different track lengths (15 or

30 m) in the five different test locations. We cannot control for this in our model as it was not noted. Furthermore, the opportunity for participants to rest during the six minutes, could also produce a bias as resting in a chair, leaning against the wall or merely standing could be measured differently by the PAMs. As we do not have the data to distinguish between and investigate these possible types of error further, we cannot investigate the magnitude or direction of this possible systematic error.

Lastly, this study did not investigate intra-model test-retest reliability, but in terms of methodology, this type of reliability is almost impossible to investigate in PAMs as the same walking pattern and hence the individual participant cannot be repeated completely. However, despite the within-individual variation in gait pattern, it would be beneficial to do an intra-person reliability test-retest study of physical activity monitors in the future.

This study also holds several strengths. To our knowledge, this study includes the largest sample size reported in the literature on validation of consumer-grade PAMs in older adults. Furthermore, this is the first study that reports results on three different models, in two different positions, and it is with another study the only one with results on subgroups using different assistive devices [12, 14–20]. The latter makes the results of this study relevant to all populations that include both older adults with and without assistive devices. The results of this validation study are easily interpreted and the conclusion should be easily transferred to research groups planning to conduct clinical studies with PAMs as an outcome measure in older adults with different use of assistive devices.

Consumer-grade PAMs can potentially replace more expensive research-grade PAMs in situations where the level of physical activity should be measured or enhanced in older adults [20]. PAMs need not have excellent validity and reliability to serve as facilitators, but if they are to be used in research settings and serve as outcome measurements, validity and reliability are key to trust the results. Clinical studies that use consumer-grade PAMs as outcome measures should use hip-worn devices, especially if the target group holds older adults with and without rollators.

Conclusion

Three of the four included consumer-grade PAMs were analysed and they showed varying measurement properties related to criterion validity among older adults performing a self-paced walking task. Our results show that wrist-worn PAMs cannot measure the number of steps in a population of older adults using rollators. The hip-worn PAMs were not significantly different in terms of measurement error or criterion validity, but when selecting a PAM for a clinical study, investigators should consider both the criterion validity and the rate of data loss as this also varied between monitors.

Abbreviations

ICC2,1: Interclass correlation coefficient 2,1; PAMs: Physical activity monitors

Acknowledgements

The authors would like to thank physiotherapist Raluca Simona Suciu for helping with the data collection, and to thank the activity centers of the municipality of Copenhagen for letting us use their facilities.

Authors' contributions

RTL contributed in setting up the project, with collecting the data, writing the manuscript and analyzing the data. CBK contributed with collecting the data and writing the manuscript. CBJ contributed in analyzing the data and writing the manuscript. HBA, HL and JC contributed in setting up the project and writing the manuscript.

Funding

The content presented within this paper was produced as part of the project REACH: this project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 690425 [28].

Availability of data and materials

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

Ethics approval and consent to participate

Oral and written information was given before participants gave informed consent to participate. The study was approved by the Danish Ethics Committee (Journal nr.:H-17033310).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Public Health, Faculty of Health and Medical Sciences, CopenRehab, Section of Social Medicine, University of Copenhagen, Gothersgade 160, 3rd floor, 1123 Copenhagen K, Denmark. [2]Research Unit of Musculoskeletal Function and Physiotherapy, Institute of Sports Science and Clinical Biomechanics, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark. [3]Department of Physiotherapy and Occupational Therapy, Copenhagen University Hospital, Herlev and Gentofte, Hellerup, Denmark. [4]Technical University of Denmark, DTU Management Engineering Inst, Diplomvej 372, 2800 Lyngby, Denmark. [5]Department of Occupational- and Physiotherapy, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark. [6]National Centre for Rehabilitation and Palliative Care, University of Southern Denmark and Odense University Hospital, Nyborg, Denmark.

Received: 22 May 2019 Accepted: 29 December 2019

References

1. WHO. WHO | Physical Activity and Older Adults. WHO. http://www.who.int/dietphysicalactivity/factsheet_olderadults/en/. Published 2011. Accessed December 18, 2017.
2. Paterson DH, Warburton DE. Physical activity and functional limitations in older adults: a systematic review related to Canada's physical activity guidelines. Int J Behav Nutr Phys Act. 2010;7:38. https://doi.org/10.1186/1479-5868-7-38.

3. Kvaavik E, Batty GD, Ursin G, Huxley R, Gale CR. Influence of individual and combined health behaviors on total and cause-specific mortality in men and women: the United Kingdom health and lifestyle survey. Arch Intern Med. 2010;170(8):711–8. https://doi.org/10.1001/archinternmed.2010.76.
4. Chodzko-zajko WJ, Proctor DN, Singh MAF, et al. Exercise and physical activity for older adults. Med Sci Sports Exerc. 2009;41(7):1510–30. https://doi.org/10.1249/MSS.0b013e3181a0c95c.
5. Szanton SL, Walker RK, Roberts L, et al. Older adults' favorite activities are resoundingly active: findings from the NHATS study. Geriatr Nur (Lond). 2015;36(2):131–5. https://doi.org/10.1016/j.gerinurse.2014.12.008.
6. Olanrewaju O, Kelly S, Cowan A, Brayne C, Lafortune L. Physical Activity in Community Dwelling Older People: A Systematic Review of Reviews of Interventions and Context. PLoS ONE. 2016;11(12). https://doi.org/10.1371/journal.pone.0168614.
7. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA. 2015;313(5):459–60. https://doi.org/10.1001/jama.2014.14781.
8. Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance the daily amount of physical activity in elderly—a protocol for a systematic review and meta-analysis. Syst Rev. 2018;7(1):69. https://doi.org/10.1186/s13643-018-0733-6.
9. Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance amount of physical activity in older adults – a systematic review and meta-analysis. Eur Rev Aging Phys Act. 2019;16(1):7. https://doi.org/10.1186/s11556-019-0213-6.
10. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737–45. https://doi.org/10.1016/j.jclinepi.2010.02.006.
11. Lee I, Park S. A comparison of gait characteristics in the elderly people, people with knee pain, and people who are Walker dependent people. J Phys Ther Sci. 2013;25(8):973–6. https://doi.org/10.1589/jpts.25.973.
12. Floegel TA, Florez-Pregonero A, Hekler EB, Buman MP. Validation of consumer-based hip and wrist activity monitors in older adults with varied ambulatory abilities. J Gerontol A Biol Sci Med Sci. 2017;72(2):229–36. https://doi.org/10.1093/gerona/glw098.
13. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act. 2015;12:159. https://doi.org/10.1186/s12966-015-0314-1.
14. Farina N, Lowry RG. The validity of consumer-level activity monitors in healthy older adults in free-living conditions. J Aging Phys Act. 2018;26(1):128–35. https://doi.org/10.1123/japa.2016-0344.
15. Paul SS, Tiedemann A, Hassett LM, et al. Validity of the Fitbit activity tracker for measuring steps in community-dwelling older adults. BMJ Open Sport Exerc Med. 2015;1(1):e000013. https://doi.org/10.1136/bmjsem-2015-000013.
16. Phillips LJ, Petroski GF, Markis NE. A comparison of accelerometer accuracy in older adults. Res Gerontol Nurs. 2015;8(5):213–9. https://doi.org/10.3928/19404921-20150429-03.
17. Thorup CB, Andreasen JJ, Sørensen EE, Grønkjær M, Dinesen BI, Hansen J. Accuracy of a step counter during treadmill and daily life walking by healthy adults and patients with cardiac disease. BMJ Open. 2017;7(3):e011742. https://doi.org/10.1136/bmjopen-2016-011742.
18. Boeselt T, Spielmanns M, Nell C, et al. Validity and usability of physical activity monitoring in patients with chronic obstructive pulmonary disease (COPD). PLoS One. 2016;11(6):e0157229. https://doi.org/10.1371/journal.pone.0157229.
19. Alharbi M, Bauman A, Neubeck L, Gallagher R. Validation of Fitbit-flex as a measure of free-living physical activity in a community-based phase III cardiac rehabilitation population. Eur J Prev Cardiol. 2016;23(14):1476–85. https://doi.org/10.1177/2047487316634883.
20. Straiton N, Alharbi M, Bauman A, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: a systematic review. Maturitas. 2018;112:85–93. https://doi.org/10.1016/j.maturitas.2018.03.016.
21. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63. https://doi.org/10.1016/j.jcm.2016.02.012.
22. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull. 1979;86(2):420–8.
23. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol. 2012;8(1):23–34.
24. Simpson L, Eng J, Klassen T, et al. Capturing step counts at slow walking speeds in older adults: comparison of ankle and waist placement of measuring device. J Rehabil Med. 2015;47(9):830–5. https://doi.org/10.2340/16501977-1993.
25. de Vet HCW. Ed. *Measurement in Medicine: A Practical Guide*. Cambridge. New York: Cambridge University Press; 2011.
26. Grant PM, Dall PM, Mitchell SL, Granat MH. Activity-monitor accuracy in measuring step number and cadence in community-dwelling older adults. J Aging Phys Act. 2008;16(2):201–14.

Publisher's Note

APPENDIX 4: PAPER D (STUDY III)

***Journal of Aging and Physical Activity*, (Ahead of Print)**
https://doi.org/10.1123/japa.2020-0214

Human Kinetics
ORIGINAL RESEARCH

Concurrent Validity Between Electronically Administered Physical Activity Questionnaires and Objectively Measured Physical Activity in Danish Community-Dwelling Older Adults

Rasmus T. Larsen, Christoffer B. Korfitsen, Carsten B. Juhl, Henning Boje Andersen, Henning Langberg, and Jan Christensen

Aim: To investigate the concurrent validity of the International Physical Activity Questionnaire-short form (IPAQ-SF) and the Nordic Physical Activity Questionnaire-short (NPAQ-short) when compared with objectively measured daily steps among older adults. ***Methods:*** Spearman's ρ between IPAQ-SF and NPAQ-short and objectively measured steps using Garmin Vivofit 3 physical activity monitors. ***Results:*** A total of 54 participants were included. The IPAQ-SF subscales' moderate physical activity (PA), moderate to vigorous PA (MVPA), and sedentary time showed little or no correlation with daily steps. The NPAQ-short subscales' vigorous PA, moderate PA, and MVPA showed little or no correlation. The IPAQ-SF subscales' vigorous PA and walking showed fair correlation. Only the IPAQ-SF metabolic equivalent of task minutes showed moderate to good correlation with daily steps. The IPAQ-SF categories and NPAQ-short categorization of World Health Organization compliance were significantly different, but the magnitudes were small and distributions indicated problems with the categorization. ***Conclusion:*** The concurrent validity is low, as the scores did not reflect objectively measured daily steps.

Keywords: IPAQ, NPAQ, physical activity monitoring

Background and Objectives

Staying physically active while aging is a key factor in a healthy lifestyle, as physically active older adults, compared with those who are inactive, have a higher physical and cognitive function, intrinsic capacity, and mobility and lower rates of pain, falls, fractures, depression, and disability (Bangsbo et al., 2019). Furthermore, inactivity among older adults is associated with an increased risk of premature all-cause mortality (Bangsbo et al., 2019). Therefore, public health recommendations and initiatives have been developed and communicated to the public to address insufficient physical activity levels. Nevertheless, 12.5% of older adults in Europe never or almost never engage in moderate or vigorous physical activity and are thus considered inactive (Gomes et al., 2017). Monitoring of physical activity levels may support health policies and public health initiatives and allow researchers to investigate patterns and changes over time (Bangsbo et al., 2019; Blair, 2009; Carnethon, 2009; Ekelund et al., 2016; Haennel & Lemire, 2002; Knuth & Hallal, 2009; Lear et al., 2017; Lee et al., 2012; Messerli, Ketelhut, & Ketelhut, 1999; Morris, Heady, Raffle, Roberts, & Parks, 1953; Wen et al., 2011).

Doubly labeled water is considered the gold standard for measuring total energy expenditure (Westerterp, 2009). However, it is rarely feasible to use total energy expenditure in large-scale studies, and consequently, accelerometry and pedometers are often used to measure physical activity objectively (Doherty et al., 2017; Pedišić & Bauman, 2015). Accelerometer-based measurement of moderate to vigorous physical activity (MVPA), with research-grade monitors, is preferred when feasible, but includes challenges with device selection and placement, measurement error, and compliance from participants, including variability in wear time (Karas et al., 2019; Westerterp, 2009). Quantifying daily steps using pedometers is a relevant alternative to accelerometry, as both measures have been shown to be associated with important outcomes, such as mortality (Lee et al., 2019; Tudor-Locke et al., 2013). Consumer-available wearables, often measuring steps by accelerometry, have shown acceptable validity when compared with visually counted number of steps (Larsen et al., 2020). They are an applicable and feasible outcome measure in health behavior research in older adults (Lyons, Swartz, Lewis, Martinez, & Jennings, 2017; Tedesco, Barton, & O'Flynn, 2017).

An alternative method for assessing physical activity is by using participant-reported outcome measures, a widely used alternative to objectively measured methods, as this is a time- and cost-efficient way to conduct the assessments (Epstein, Miller, Stitt, & Morris, 1976; Sallis & Saelens, 2000). However, the psychometric properties, and thus the applicability, of these participant-reported physical activity measurements are challenged by several forms of bias, including recall and social desirability bias, that ultimately can result in less precise assessments (Colley, Butler, Garriguet, Prince, & Roberts, 2018; Helmerhorst, Brage, Warren, Besson, & Ekelund, 2012; Lee, Macfarlane, Lam, & Stewart, 2011; Prince et al., 2008, 2020; Sagelv et al., 2020; Silsbury, Goldsmith, &

Larsen and Korfitsen are with the Section of Social Medicine, Department of Public Health, University of Copenhagen, Copenhagen, Denmark. Korfitsen is also with Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark. Juhl is with the Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark; and the Department of Physiotherapy and Occupational Therapy, Copenhagen University Hospital, Herlev and Gentofte, Denmark. Andersen is with the Department of Technology, Management and Economics, Technical University of Denmark, Lyngby, Denmark. Langberg is with the Section for Health Services Research, Department of Public Health, University of Copenhagen, Copenhagen, Denmark. Christensen is with the Department of Occupational Therapy and Physiotherapy, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark; and the National Centre for Rehabilitation and Palliative Care, University of Southern Denmark, Copenhagen, Denmark; and Odense University Hospital, Odense, Denmark. Larsen (rala@sund.ku.dk) is corresponding author.

Rushton, 2015). Because of this, the investigation of concurrent validity of participant-reported physical activity measurements and thus the ability to reflect objectively measured physical activity, including daily steps, provides knowledge of the applicability of the measurements, which in turn, is crucial for guiding researchers and clinicians in designing studies and interpreting results.

Several participant-reported physical activity measurements have been developed, with slightly different purposes. Among these are the International Physical Activity Questionnaire-short form (IPAQ-SF) and the Nordic Physical Activity Questionnaire-short (NPAQ-short). IPAQ-SF is widely used in research projects and in population-level assessments and is utilized as an easy tool to evaluate self-reported MVPA, walking activity, and sedentary behavior (Cleland, Ferguson, Ellis, & Hunter, 2018; Craig et al., 2003; Lee et al., 2011; Loyen et al., 2016). NPAQ-short was developed to be utilized on a population level by the Danish National Health Profile in 2017 (Danish Health Authority, 2018; Danquah, Petersen, Skov, & Tolstrup, 2018). Nevertheless, the psychometric properties, including the criterion and concurrent validity of IPAQ-SF and NPAQ-short among older adults, is unclear (Danquah et al., 2018; Lee et al., 2011; Silsbury et al., 2015).

Thus, the aim of this study was to investigate the concurrent validity of electronically and self-administered IPAQ-SF and NPAQ-short when compared with objectively measured daily steps, and, subsequently, to investigate response patterns to assess the applicability of IPAQ-SF and NPAQ-short among community-dwelling older adults.

Methods

This study was conducted using baseline data from the Motivational Interviewing and Physical Activity Monitoring (MIPAM) randomized controlled trial to enhance the daily level of physical activity among older adults. The MIPAM trial and this study were conducted with waivers from the Danish Ethics Committee of the Capital Region of Denmark (journal number: 18004960). The Danish Data Protection Agency approved the handling of data (reference number: 514-0268/18-3000). Informed consent from the participants was collected electronically via SurveyXact (Copenhagen, Denmark). Prior to agreeing and signing the consent survey, the participants received written information about the study.

Procedures and Measures

Before randomization, the participants received the hip-worn physical activity monitor (PAM; Garmin Vivofit 3; Garmin International, Inc., Olathe, KS) by e-mail and were asked to wear it for a week, which would become the baseline week. The participants received the necessary instruction and support in using the PAM and the application. After the baseline week, the participants received an electronic survey containing the participant-reported physical activity questionnaires.

Participants

Older adults were recruited through social media and with advertisements in local activity centers. Participants were considered eligible for inclusion in the MIPAM trial if they (a) were retired and community dwelling, (b) were at least 70 years of age by the day of enrollment in the trial, (c) owned a smartphone or tablet and were able to install the Garmin Connect application (Olathe, KS), (d) had an active e-mail account and were able to complete the electronic study survey, and (e) had hearing ability sufficient to participate in a telephone interview. Participants were excluded if they suffered from a moderate or severe decline in mental ability (e.g., from Alzheimer's disease), were undergoing active chemotherapy or palliative care for cancer, or had a major mobility impairment preventing them from walking. Furthermore, to be considered eligible for inclusion in the present study, the participants had to provide daily step data for at least 4 days and should have completed the electronic survey.

Daily Steps as Objectively Measured Physical Activity

In a sample of older adults, our research group has investigated the criterion validity of four commercially available PAMs and found that the hip-worn Garmin Vivofit 3 showed acceptable validity in relation to counting steps (Larsen et al., 2020). Hence, in this study, daily steps obtained from the PAM is considered the construct of interest, using the average number of daily steps during the baseline week.

Physical Activity Questionnaires

International Physical Activity Questionnaire-Short Form. The IPAQ-SF consists of seven items and assesses the amount of MVPA, metabolic equivalent of task minutes (MET-minutes), walking time, and sedentary time, with a recall frame of "the last 7 days." From the IPAQ responses, three physical activity levels can be derived: "low activity level," "moderate activity level," and "high activity level" (Craig et al., 2003). An international validation study of the IPAQ-SF assessment has reported acceptable test–retest reliability and recommends the use of the "last 7 days" version of the measurement (Craig et al., 2003). However, a more recent systematic review reports low criterion validity between IPAQ-SF and objectively measured physical activity (Lee et al., 2011).

Nordic Physical Activity Questionnaire-Short. The two-item NPAQ-short (Danquah et al., 2018) is a short revised version of the original NPAQ. It was developed to monitor MVPA time and compliance with the World Health Organization (WHO) recommendations on physical activity in a typical week in population-based surveys (Danquah et al., 2018; World Health Organization, 2010). NPAQ-short has shown a fair correlation with objectively measured MVPA (Spearman's $\rho = .33$) in a Danish adult population (Danquah et al., 2018). From NPAQ-short, the following four physical activity categories can be calculated: "inactive," "not sufficiently physically active," "sufficiently physically active," and "optimally physically active." NPAQ-short's definition of compliance with WHO's recommendations included sufficiently and optimally physically active participants.

Statistical Analyses

Distributions of continuous data were evaluated by quantile–quantile plots and histograms of the standardized residuals. For a description of measurement response patterns, continuous data from the measurements were summarized, using both parametric and nonparametric statistics. Categorical data were summarized with frequencies and percentages of the total score. The number of participants reporting zero activity were summarized as binary variables. Spearman's rank correlation coefficients between the measurement scores and daily steps were calculated. We used the following interpretation of Spearman's rank correlation coefficients for concurrent validity: 0–.25 = little or no correlation; .25–.50 = fair correlation; .50–.75 = moderate to good correlation; and above .75 = good to excellent correlation (Portney & Watkins, 2014, p. 535). A priori, we hypothesized the IPAQ-SF and NPAQ-short scores to have at least a fair correlation with daily steps. However, for "IPAQ-SF Walking" and "NPAQ-short Moderate Physical Activity," we expected to find a moderate to good correlation with daily steps as the description of the specific questions including walking. Finally, for "IPAQ-SF Moderate Activity," we expected to find little or no correlation with daily steps as the description of the specific question specifically excludes walking activities (Portney & Watkins, 2014, p. 535). For IPAQ-SF categories and NPAQ-short categories, we investigated the relationship with one-way analysis of variance models and unpaired *t* tests. The results of these were presented as means and 95% confidence intervals (95% CIs) and were visualized by scatterplots of instrument scores and daily steps and a bar chart of correlation scores between instrument scores and daily steps. The IPAQ-SF can be categorized into three activity levels. The low activity level includes participants who did not meet the criteria for moderate- and vigorous-intensity categories. The moderate activity level includes participants who had (a) 3 or more days of vigorous-intensity activity of at least 20 min/day, (b) 5 or more days of moderate-intensity activity and/or walking of at least 30 min/day, or (c) 5 or more days of any combination of walking and moderate- or vigorous-intensity activities, achieving a minimum total physical activity of at least 600 MET-min/week. The high activity level includes participants who had (a) vigorous-intensity activity on at least 3 days (20 min minimum, achieving a minimum of total physical activity of at least 1,500 MET-min/week) or (b) 7 or more days of any combination of walking and moderate- or vigorous-intensity activities achieving a minimum total physical activity of at least 3,000 MET-min/week. NPAQ-short is likewise categorized into four activity categories: inactive participants, not sufficiently physically active participants, sufficiently physically active participants, and optimally physically active participants.[1]

RStudio (version 1.1.463; 2016, Integrated Development for R; RStudio, Inc., Boston, MA) was used for all statistical analyses, and visualizations were created with the "ggplot2" package. An alpha level of .05 was considered the threshold for statistical significance.

Results

A total of 67 eligible participants answered the baseline survey. Of these, 13 were excluded for having fewer than 4 days of objectively measured steps per day during the baseline week, leaving 54 participants included in the study. Thirty (56%) of the included participants were women, the median age was 71.5 years, and the median self-reported percentage of physical activity that consisted of walking was reported to be 70.0%. Table 1 reports the characteristics of the 54 included participants, number of daily steps, and IPAQ-SF and NPAQ-short scores. Supplementary Figures S1 and S2 (available online) illustrate the distributions of the IPAQ-SF and NPAQ-short scores, as well as the association with daily steps for the participants, using scatterplots, best-fitted lines, and CIs.

Table 2 reports the concurrent validity between daily steps and the IPAQ-SF and NPAQ-short scores. Figure 1 presents the correlation coefficients for concurrent validity (presented in Table 2) and lines for the interpretation of Spearman's rank correlation coefficients. IPAQ-SF moderate-intensity minutes, MVPA time, and sedentary time had little or no correlation with daily steps. The NPAQ-short scores with little or no correlation include vigorous activity, moderate activity, MVPA, and compliance with the WHO's recommendations. The IPAQ-SF scores with fair correlation include vigorous activity and walking. Only the IPAQ-SF MET-minutes showed moderate to good correlation with daily steps. Only the IPAQ-SF vigorous activity, MET-minutes, and walking had a significant correlation with daily steps.

For the categorical constructs of the questionnaires, one-way analysis of variance revealed a difference in daily steps between the IPAQ-SF categories ($p < .001$). The difference between the means in the low and high categories was −3,531 steps/day (95% CI [−5,497, −1,565]), the difference between the means in the low and moderate categories was −2,561 steps/day (95% CI [−4,970, −332]), and the difference between the means in the moderate and high categories was −970 steps/day (95% CI [−3,410, 1,471]). One-way analysis of variance revealed no differences in daily steps between the four NPAQ-short categories ($p = .24$). An unpaired *t* test revealed that the mean daily steps of the two categories in the NPAQ-short compliance with WHO recommendations were different ($p = .04$). The participants who, according to the NPAQ-short, were categorized to comply with the WHO recommendations walked on average 1,628 steps/day (95% CI [76, 3,180]) more than the participants who were categorized not to comply. The distributions, means, and 95% CI are illustrated in Figure 2.

Discussion

To our knowledge, this is the first study to investigate the ability of electronically administered IPAQ-SF and NPAQ-short to reflect objectively measured daily steps among Danish older adults. In this study, the electronically administered IPAQ-SF and NPAQ-short did not adequately reflect objectively measured daily steps and, thus, is not recommended to be used to predict physical activity levels among older adults. Compared with previous studies on older adults for the IPAQ-SF and adults for the NPAQ-short, we found lower correlation coefficients for validity (Danquah et al., 2018; Deng et al., 2008; Kolbe-Alexander, Lambert, Harkins, & Ekelund, 2006). This might be explained by the fact that the previous IPAQ-SF and NPAQ-short results were obtained in younger samples (mean age of 65, 67, and 43 years, respectively) and/or as a consequence of IPAQ-SF being administered by interview (Danquah et al., 2018; Deng et al., 2008; Kolbe-Alexander et al., 2006). However, when interpreting our results, it should be taken into account that the NPAQ-short study used accelerometer counts and, thus, MVPA, as the construct of interest (Danquah et al., 2018), which may explain why the authors found results that differ from those of the present study, which used daily steps. Nevertheless, the IPAQ-SF seems to perform slightly better than the NPAQ-short, especially when comparing the IPAQ-SF MET and IPAQ-SF Walking with the findings from the NPAQ-short. This might be explained by several factors. First, IPAQ-SF recalls the past 7 days of the participants, whereas the NPAQ-short recalls a typical week. So, if the participants were affected by the Hawthorne effect and thus changed their behavior when enrolled into the study, this might have caused some bias in the results. On the other hand, previous findings have shown that two different versions of the IPAQ-SF using two different recall frames, "usual week" and "last 7 days," perform similarly, even though the recall bias might be introduced using recall participant-reported physical activity measurements. Second, identifying a criterion of interest

Table 1 Demographics, Daily Steps, and Instrument Scores

Variable	Total sample ($N = 54$)		
Gender			
Women, *n* (%)	30 (55.6%)		
Walking aid			
None, *n* (%)	53 (98.1%)		
Cane, *n* (%)	1 (1.9%)		
	Mean (*SD*)	**Median (IQR)**	
Age (years)	73.1 (3.9)	71.5 (70–74.0)	
Height (cm)	171 (9.5)	171.5 (165.0–177.8)	
Weight (kg)	80.3 (14.9)	79.0 (70.3–86.8)	
Body mass index (kg/m^2)	27.4 (4.4)	26.8 (24.3–29.1)	
Walking as percent of total reported PA	58.4% (31.3%)	70.0% (40.0%–80.0%)	
Daily steps	5,782.4 (3,005.4)	5,682.9 (3,422.4–8,036.1)	
	Mean (*SD*)	**Median (IQR)**	**Frequency of zero activity, *n*(%)**
IPAQ-SF			
Vigorous PA (min/week)	144.7 (390.9)	0.0 (0.0–141.3)	30 (56%)
Moderate PA (min/week)	185.4 (418.0)	25.0 (0.0–232.5)	25 (46%)
MET (min/week)	2,972.1 (5,126.5)	1,410 (594.0–3,318.8)	6 (11%)
MVPA (min/week)	330.0 (781.1)	62.5 (0.0–350.0)	18 (33%)
Sedentary (min/day)	329.4 (184.0)	300.0 (192.5–465.0)	2 (4%)
Walking (min/week)	325.3 (465.0)	180.0 (0.00–420.0)	20 (37%)
PA categories			
Low activity level, *n* (%)	27 (50.0%)		
Moderate activity level, *n* (%)	11 (20.4%)		
High activity level, *n* (%)	16 (29.6%)		
	Mean (*SD*)	**Median (IQR)**	**Frequency of zero activity, *n*(%)**
NPAQ-short			
Vigorous PA (min/week)	142.0 (183.7)	79.5 (0.00–172.5)	18 (33%)
Moderate PA (min/week)	269.7 (399.7)	120.0 (0.0–286.8)	15 (28%)
MVPA (min/week)	411.6 (447.4)	240.0 (122.5–592.5)	6 (11%)
PA categories			
Inactive	6 (11.1%)		
Not sufficiently physically active	7 (13.0%)		
Sufficiently physically active	10 (18.5%)		
Optimally physically active	31 (57.4%)		
NPAQ-short compliance with WHO's recommendations			
No	13 (24.1%)		
Yes	41 (75.9%)		

Note. Daily steps per day and height were considered as normally distributed. All other continuous variables were judged as having a nonnormal distribution. Number of "zero" responses is defined as participants who reported no activity in a specific category. IPAQ-SF categories: Low activity level refers to participants who did not meet the criteria for moderate- and vigorous-intensity categories. Moderate activity level refers to participants who had (a) 3 or more days of vigorous-intensity activity of at least 20 min/day, (b) 5 or more days of moderate-intensity activity and/or walking of at least 30 min/day, or (c) 5 or more days of any combination of walking and moderate- or vigorous-intensity activities achieving a minimum total PA of at least 600 MET-min/week. High activity level refers to participants who had (a) vigorous-intensity activity on at least 3 days (20 min minimum, achieving a minimum total PA of at least 1,500 MET-min/week) or (b) 7 or more days of any combination of walking and moderate- or vigorous-intensity activities achieving a minimum total PA of at least 3,000 MET-min/week. NPAQ-short categories: Inactive participants were defined as participants with no moderate or vigorous PA, not sufficiently physically active (moderate PA/150 + vigorous PA/75) < 1.0 and (moderate PA/150 + vigorous PA/75) > 0, sufficiently physically active (moderate PA/150 + vigorous PA/75) ≥ 1.0 and (moderate PA/150 + vigorous PA/75) < 2.0, and optimally physically active (moderate PA/150 + vigorous PA/75) ≥ 2.0. Compliance with WHO's recommendations included sufficiently and optimally physically active participants. IQR = interquartile range; PA = physical activity; IPAQ-SF = International Physical Activity Questionnaire-short form; MET = metabolic equivalent of task; MVPA = moderate to vigorous physical activity; NPAQ-short = Nordic Physical Activity Questionnaire-short; WHO = World Health Organization.

Table 2 Concurrent Validity as Spearman's Rank Correlation Coefficients Between Instrument Scores and Objectively Measured Daily Steps

Physical activity questionnaire	Correlation as Spearman's ρ	*p*	Interpreted correlation	A priori expected correlation
IPAQ-SF				
Vigorous PA (min/week)	.29	.04	Fair	Fair
Moderate PA (min/week)	.12	.41	Little or no	Little or no
MVPA (min/week)	.24	.08	Little or no	Fair
MET (min/week)	.50	<.01	Moderate to good	Fair
Sedentary (min/day)	−.08	.59	Little or no	Fair
Walking (min/week)	.45	<.01	Fair	Moderate to good
NPAQ-short				
Vigorous PA (min/week)	.12	.40	Little or no	Fair
Moderate PA (min/week)	.15	.29	Little or no	Moderate to good
MVPA (min/week)	.22	.11	Little or no	Fair

Note. Interpretation of Spearman's rank correlation coefficients for concurrent validity: 0–.25 = little or no correlation; .25–.50 fair correlation; .50–.75 = moderate to good correlation; and above .75 = good to excellent correlation. IPAQ-SF = International Physical Activity Questionnaire-short form; PA = physical activity; MVPA = moderate to vigorous physical activity; MET = metabolic equivalent of task; NPAQ-short = Nordic Physical Activity Questionnaire-short.

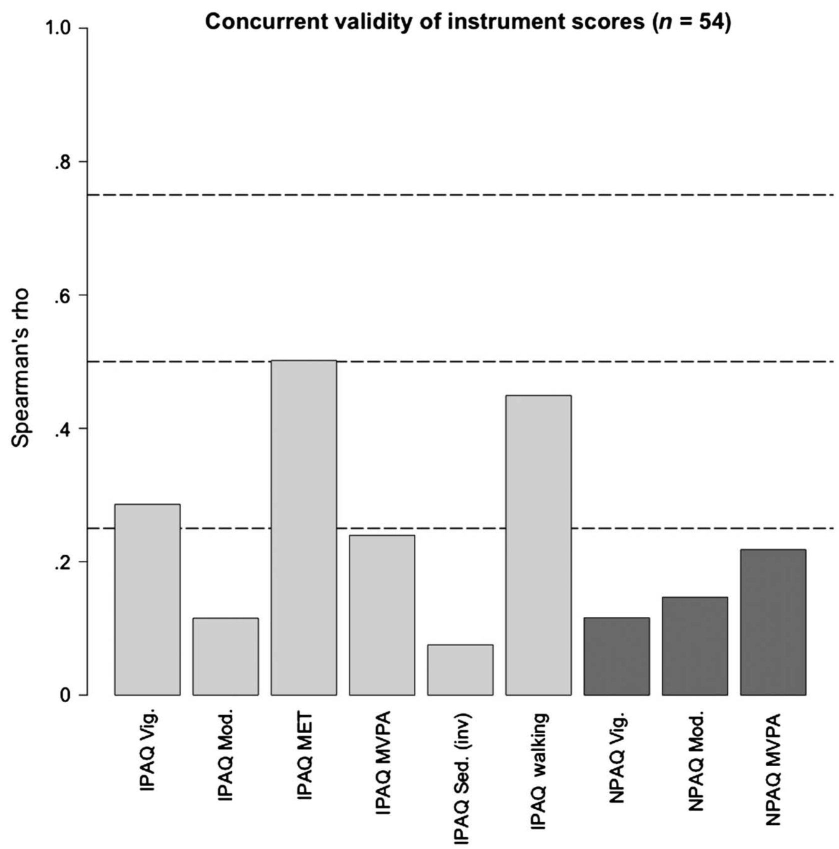


Figure 1 — Performance of instrument scores expressed as the correlation between the instrument and objectively measured steps per day. Light gray-colored bars represent IPAQ-SF constructs, and dark gray-colored bars represent NPAQ-short constructs. *y*-axis: Spearman's rank correlation coefficient (ρ). Reference lines: below the lowest line indicates little or no correlation, above the lowest line indicates a fair correlation, above the middle line indicates a moderate to good correlation, and above the top line indicates a good to excellent correlation. IPAQ-SF sedentary construct is inverse, meaning the true value is the negative equivalent of the illustrated. IPAQ = International Physical Activity Questionnaire; Vig. = vigorous; Mod. = moderate; MET = metabolic equivalent of task; MVPA = moderate to vigorous physical activity; SF = short form; Sed. = sedentary; NPAQ = Nordic Physical Activity Questionnaire.

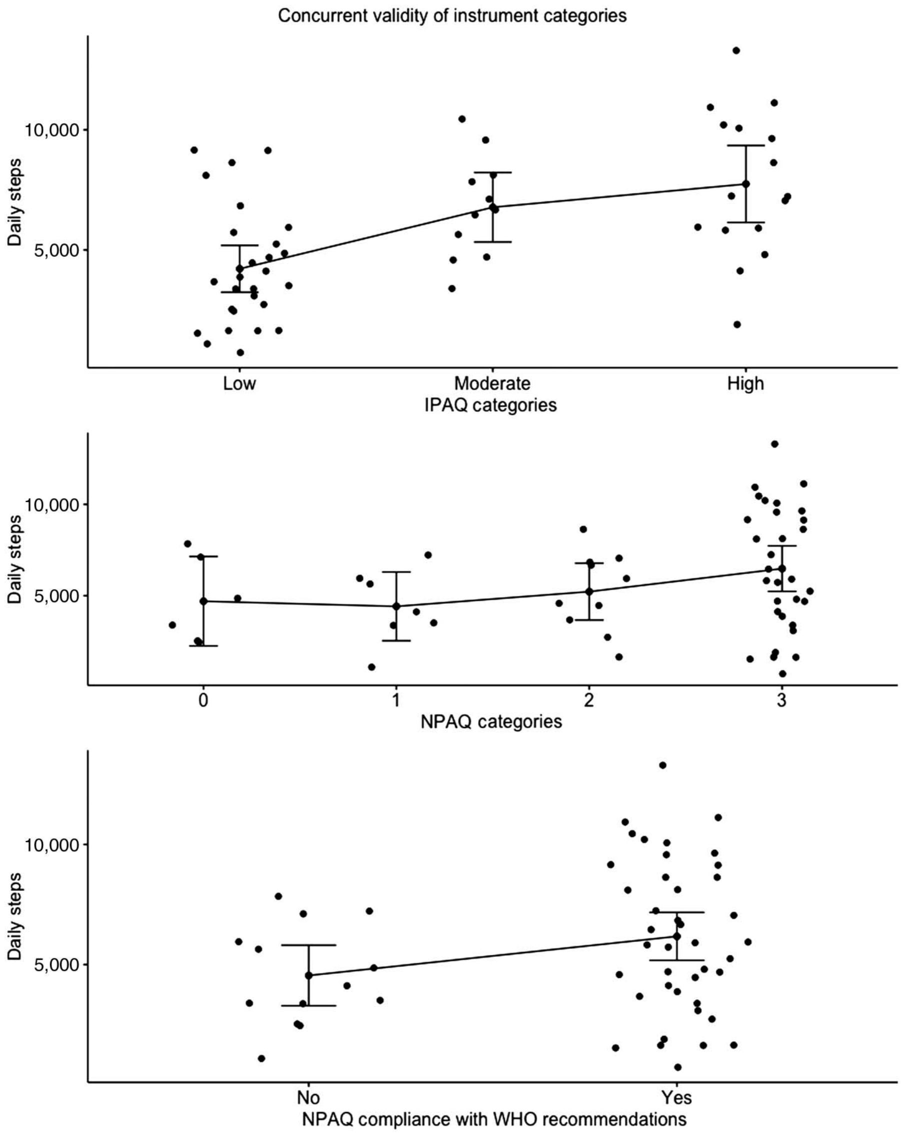


Figure 2 — Means, confidence intervals, and distributions of average steps per day of the IPAQ-short form categorical construct and the NPAQ-short categorical constructs. NPAQ-short categories: (a) inactive participants were defined as participants with no moderate or vigorous physical activity, (b) not sufficiently physically active (moderate physical activity/150 + vigorous physical activity/75) < 1.0 and (moderate physical activity/150 + vigorous physical activity/75) > 0, (c) sufficiently physically active (moderate physical activity/150 + vigorous physical activity/75) ≥ 1.0 and (moderate physical activity/150 + vigorous physical activity/75) < 2.0, and (d) optimally physically active (moderate physical activity/150 + vigorous physical activity/75) ≥ 2.0. Compliance with WHO's recommendations included sufficiently and optimally physically active participants. IPAQ = International Physical Activity Questionnaire; NPAQ = Nordic Physical Activity Questionnaire; WHO = World Health Organization.

that reflects a multidimensional construct such as physical activity is a difficult task, as multiple factors could be considered (Sylvia, Bernstein, Hubbard, Keating, & Anderson, 2014). Research-grade accelerometers have been found to measure MET-minutes and MVPA time with acceptable validity (Prince et al., 2008), which is considered the gold standard of physical activity measurement (Westerterp, 2009). However, steps per day can similarly serve as a relevant construct of physical activity in older adults for many reasons. First, older adults rate walking as their favorite physical activity (followed by gardening and playing sports; Szanton et al., 2015), which is supported by our results of self-reported walking at 70%. Second, the recommendations to older adults of relevant physical activity by the WHO, the American College of Sports Medicine, and the American Heart Association mainly consist of stepping activities (Elsawy & Higgins, 2010; Nelson et al., 2007; World Health Organization, 2011). Finally, emerging evidence suggests that the number of daily steps, not the intensity of the walking activity, is a predictor for all-cause mortality in older adults (Lee et al., 2019). This underpins the proposition that the number of daily steps is a reasonable criterion measure of physical activity in older adults and is practically applicable via easy and low-cost capture by consumer-grade PAMs (Larsen et al., 2020). However, the term "criterion validity" was not chosen to be used in this study because of the multidimensionality of the physical activity questionnaires. Instead, we conservatively framed this study to investigate concurrent validity of the physical activity questionnaires and whether they reflected objectively measured daily steps, a highly relevant construct of physical activity among older adults.

The individual response patterns of both the IPAQ-SF and the NPAQ-short show that participant-reported outcome measures tend to follow response patterns (i.e., floor effect) that challenge the feasibility of using the IPAQ-SF and the NPAQ-short for older adults. Thirty-seven percent of the participants reported no walking activities exceeding 10 min. This is also relevant to the NPAQ-short, as 28% of the participants reported no moderate physical activity, a construct that specifically includes stepping activities such as walking. This is illustrated by histograms in the Supplementary Figures S1 and S2 (available online). The discrepancy between daily steps and the "zero" response in the physical activity measurements indicates that some of the participants included in this study did not understand the examples in the physical activity measurements or had issues with the reporting or the administration. Furthermore, the results from the analyses of the categorical constructs and the illustrations in Figure 2 also highlight the problem. Even though the one-way analysis of variance of the IPAQ-SF categories and the unpaired *t* test of the NPAQ-short dichotomic construct revealed significant differences among the sample means, the magnitudes of the differences and the distributions from Figure 2 indicate a problem with the questionnaires, as the categories fail to differentiate sufficiently between the participants. Findings from cognitive interviews of older community-dwelling adults suggest that there is a need for caution in administering the IPAQ among older adults, as difficulties occurred in most stages of answering the questions and errors resulted in both under- and overreporting (Heesch, van Uffelen, Hill, & Brown, 2010). The reason for the "zero" response in our study is unknown. This might be related to the findings from the cognitive interviews (Heesch et al., 2010), but could also be related to the electronic administration of the measurements. Whatever the reason for the "zero" responses, it leads to the question of whether participant-reported physical activity measurements should be used among older adults. More detailed descriptions in the measurements, an adapted and validated version of the IPAQ specifically for older adults (Rubio Castañeda, Tomás Aznar, & Muro Baquero, 2017), and alternatives to administering the forms electronically should be considered (Cleland et al., 2018).

Strength and Limitations

This is the first study to investigate the ability of the two electronically administered physical activity questionnaires, the IPAQ-SF and the NPAQ-short, to validly reflect objectively measured daily steps. As mentioned earlier, daily steps serves as a highly relevant physical activity construct among older adults (Hansen et al., 2020; Lee et al., 2019). However, the physical activity questionnaires do not estimate step counts. Because of this, this study was framed as a concurrent validity study instead of a criterion validity study (Mokkink, Terwee, Patrick, et al., 2010; Portney & Watkins, 2014).

Despite a limited sample size of 54 participants, the sample size is considered to be "very good" to investigate concurrent validity (Mokkink, Terwee, Knol, et al., 2010). Furthermore, our findings are in line with previous research on IPAQ-SF (Danquah et al., 2018; Lee et al., 2011; Silsbury et al., 2015), and we believe that our results indicate that caution must be exercised in using these self-reported physical activity questionnaires by electronic administration for community-dwelling older adults. We have identified a problematic response pattern for both questionnaires, but cannot identify the sources of these in the way of response administration, nor in the study setup. Therefore, future research should investigate response patterns to assess whether self-reported physical activity measurements can be applied to older adults. Furthermore, for emphasizing specific walking activities that otherwise might be ignored by older adults, examples such as walking to the grocery store, walking the dog, or walking for exercise should be listed in the physical activity questionnaires. The generalizability of our results might be threatened by our inclusion criteria, that participants had to own a smartphone and a valid e-mail address, and thereby represent a selected group. The general population of Danish older adults might have even lower technological competencies and the way of administering the questionnaires might cause even larger problems in the background population. Finally, as Figure 2 illustrates, the physical activity categories did not adequately distinguish between activity levels of the participants. However, this finding might have changed if we included a large sample with more inactive participants, and thus, the risk of having an underrepresentation of inactive participants could exist. Nonetheless, as the mean daily step count was similar to the median daily step count, the distribution of daily steps was not skewed in either direction, and the abovementioned consideration did, therefore, not affect our results directly.

Conclusions

The concurrent validity was found to be inadequate, as the electronically administered IPAQ-SF and NPAQ-short did not reflect objectively measured daily steps. There was a lack of understanding of the physical activity questionnaires among the participants, as an investigation of the response patterns revealed that a considerable number of the participants reported "zero" physical activity despite the objective data suggesting otherwise. The IPAQ-SF vigorous score and the IPAQ-SF MET-minutes fulfilled our prior hypothesis of a fair concurrent validity. The IPAQ-SF walking construct also had a fair concurrent validity, but was expected to perform better, with at least a moderate to good criterion validity. The results from electronically administered IPAQ-SF and NPAQ-short should be interpreted with caution due to low concurrent validity and inconsistency when reporting physical activity.

Note

1. Inactive participants were defined as participants with no moderate or vigorous physical activity, not sufficiently physically active participants (moderate physical activity/150 + vigorous physical activity/75) < 1.0 and (moderate physical activity/150 + vigorous physical activity/75) > 0, sufficiently physically active participants (moderate physical activity/150 + vigorous physical activity/75) ≥ 1.0 and (moderate physical activity/150 + vigorous physical activity/75) < 2.0, and optimally physically active participants (moderate physical activity/150 + vigorous physical activity/75) ≥ 2.0.

References

Bangsbo, J., Blackwell, J., Boraxbekk, C.-J., Caserotti, P., Dela, F., Evans, A.B., … Viña, J. (2019). Copenhagen consensus statement 2019: Physical activity and ageing. *British Journal of Sports Medicine, 53*(14), 856–858. PubMed ID: 30792257 doi:10.1136/bjsports-2018-100451

Blair, S.N. (2009). Physical inactivity: The biggest public health problem of the 21st century. *British Journal of Sports Medicine, 43*(1), 1–2. PubMed ID: 19136507

Carnethon, M.R. (2009). Physical activity and cardiovascular disease: How much is enough? *American Journal of Lifestyle Medicine, 3*(1, Suppl.), 44S–49S. doi:10.1177/1559827609332737

Cleland, C., Ferguson, S., Ellis, G., & Hunter, R.F. (2018). Validity of the International Physical Activity Questionnaire (IPAQ) for assessing moderate-to-vigorous physical activity and sedentary behaviour of older adults in the United Kingdom. *BMC Medical Research Methodology, 18*(1), 176. PubMed ID: 30577770 doi:10.1186/s12874-018-0642-3

Colley, R.C., Butler, G., Garriguet, D., Prince, S.A., & Roberts, K.C. (2018). Comparison of self-reported and accelerometer-measured physical activity in Canadian adults. *Health Reports, 29*(12), 3–15. PubMed ID: 30566204

Craig, C.L., Marshall, A.L., Sjöström, M., Bauman, A.E., Booth, M.L., Ainsworth, B.E., . . . Oja, P. (2003). International Physical Activity Questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise, 35*(8), 1381–1395. PubMed ID: 12900694 doi:10.1249/01.MSS.0000078924.61453.FB

Danish Health Authority. (2018). *The health of the Danes—The National Health Profile 2017* [Danskernes Sundhed—Den Nationale Sundhedsprofil 2017]. Copenhagen, Denmark: Sundhedsstyrelsen.

Danquah, I.H., Petersen, C.B., Skov, S.S., & Tolstrup, J.S. (2018). Validation of the NPAQ-short—A brief questionnaire to monitor physical activity and compliance with the WHO recommendations. *BMC Public Health, 18*(1), 601. PubMed ID: 29739383 doi:10.1186/s12889-018-5538-y

Deng, H.B., Macfarlane, D.J., Thomas, G.N., Lao, X.Q., Jiang, C.Q., Cheng, K.K., & Lam, T.H. (2008). Reliability and validity of the IPAQ-Chinese: The Guangzhou Biobank Cohort study. *Medicine & Science in Sports & Exercise, 40*(2), 303–307. PubMed ID: 18202571 doi:10.1249/mss.0b013e31815b0db5

Doherty, A., Jackson, D., Hammerla, N., Plötz, T., Olivier, P., Granat, M.H., . . . Wareham, N.J. (2017). Large scale population assessment of physical activity using wrist worn accelerometers: The UK Biobank study. *PLoS ONE, 12*(2), e0169649. PubMed ID: 28146576 doi:10.1371/journal.pone.0169649

Ekelund, U., Steene-Johannessen, J., Brown, W.J., Fagerland, M.W., Owen, N., Powell, K.E., . . . Lee, I.-M. (2016). Does physical activity attenuate, or even eliminate, the detrimental association of sitting time with mortality? A harmonised meta-analysis of data from more than 1 million men and women. *The Lancet, 388*(10051), 1302–1310. doi:10.1016/S0140-6736(16)30370-1

Elsawy, B., & Higgins, K.E. (2010). Physical activity guidelines for older adults. *American Family Physician, 81*(1), 55–59. PubMed ID: 20052963

Epstein, L., Miller, G.J., Stitt, F.W., & Morris, J.N. (1976). Vigorous exercise in leisure time, coronary risk-factors, and resting electrocardiogram in middle-aged male civil servants. *British Heart Journal, 38*(4), 403–409. PubMed ID: 1267984 doi:10.1136/hrt.38.4.403

Gomes, M., Figueiredo, D., Teixeira, L., Poveda, V., Paúl, C., Santos-Silva, A., & Costa, E. (2017). Physical inactivity among older adults across Europe based on the SHARE database. *Age and Ageing, 46*(1), 71–77. PubMed ID: 28181637 doi:10.1093/ageing/afw165

Haennel, R.G., & Lemire, F. (2002). Physical activity to prevent cardiovascular disease. How much is enough? *Canadian Family Physician Medecin De Famille Canadien, 48,* 65–71. PubMed ID: 11852614

Hansen, B.H., Dalene, K.E., Ekelund, U., Fagerland, M.W., Kolle, E., Steene-Johannessen, J., . . . Anderssen, S.A. (2020). Step by step: Association of device-measured daily steps with all cause-mortality – A prospective cohort study. *Scandinavian Journal of Medicine & Science in Sports, 30*(9), 1705–1711. PubMed ID: 32427398 doi:10.1111/sms.13726

Heesch, K.C., van Uffelen, J.G., Hill, R.L., & Brown, W.J. (2010). What do IPAQ questions mean to older adults? Lessons from cognitive interviews. *International Journal of Behavioral Nutrition and Physical Activity, 7*(1), 35. PubMed ID: 20459758 doi:10.1186/1479-5868-7-35

Helmerhorst, H.J.F., Brage, S., Warren, J., Besson, H., & Ekelund, U. (2012). A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *International Journal of Behavioral Nutrition and Physical Activity, 9*(1), 103. PubMed ID: 22938557 doi:10.1186/1479-5868-9-103

Karas, M., Bai, J., Strączkiewicz, M., Harezlak, J., Glynn, N.W., Harris, T., . . . Urbanek, J.K. (2019). Accelerometry data in health research: Challenges and opportunities. *Statistics in Biosciences, 11*(2), 210–237. PubMed ID: 31762829 doi:10.1007/s12561-018-9227-2

Knuth, A.G., & Hallal, P.C. (2009). Temporal trends in physical activity: A systematic review. *Journal of Physical Activity & Health, 6*(5), 548–559. PubMed ID: 19953831 doi:10.1123/jpah.6.5.548

Kolbe-Alexander, T.L., Lambert, E.V., Harkins, J.B., & Ekelund, U. (2006). Comparison of two methods of measuring physical activity in South African older adults. *Journal of Aging and Physical Activity, 14*(1), 98–114. PubMed ID: 16648654 doi:10.1123/japa.14.1.98

Larsen, R.T., Korfitsen, C.B., Juhl, C.B., Andersen, H.B., Langberg, H., & Christensen, J. (2020). Criterion validity for step counting in four consumer-grade physical activity monitors among older adults with and without rollators. *European Review of Aging and Physical Activity, 17*(1), 1. doi:10.1186/s11556-019-0235-0

Lear, S.A., Hu, W., Rangarajan, S., Gasevic, D., Leong, D., Iqbal, R., . . . Yusuf, S. (2017). The effect of physical activity on mortality and cardiovascular disease in 130 000 people from 17 high-income, middle-income, and low-income countries: The PURE study. *The Lancet, 390*(10113), 2643–2654. doi:10.1016/S0140-6736(17)31634-3

Lee, I.-M., Shiroma, E.J., Kamada, M., Bassett, D.R., Matthews, C.E., & Buring, J.E. (2019). Association of step volume and intensity with all-cause mortality in older women. *JAMA Internal Medicine, 179*(8), 1105–1112. doi:10.1001/jamainternmed.2019.0899

Lee, I.-M., Shiroma, E.J., Lobelo, F., Puska, P., Blair, S.N., Katzmarzyk, P.T., & Lancet Physical Activity Series Working Group. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: An analysis of burden of disease and life expectancy. *The Lancet, 380*(9838), 219–229. doi:10.1016/S0140-6736(12)61031-9

Lee, P.H., Macfarlane, D.J., Lam, T.H., & Stewart, S.M. (2011). Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): A systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 8*(1), 115. PubMed ID: 22018588 doi:10.1186/1479-5868-8-115

Loyen, A., Van Hecke, L., Verloigne, M., Hendriksen, I., Lakerveld, J., Steene-Johannessen, J., . . . van der Ploeg, H.P. (2016). Variation in population levels of physical activity in European adults according to cross-European studies: A systematic literature review within DEDIPAC. *International Journal of Behavioral Nutrition and Physical Activity, 13*(1), 72. PubMed ID: 27350359 doi:10.1186/s12966-016-0398-2

Lyons, E.J., Swartz, M.C., Lewis, Z.H., Martinez, E., & Jennings, K. (2017). Feasibility and acceptability of a wearable technology physical activity intervention with telephone counseling for mid-aged and older adults: A randomized controlled pilot trial. *JMIR mHealth and uHealth, 5*(3), e28. PubMed ID: 28264796 doi:10.2196/mhealth.6967

Messerli, F.H., Ketelhut, R.G., & Ketelhut, K. (1999). Cardiovascular risk factors and physical activity: How much is enough? *Archives of*

Internal Medicine, 159(8), 882–883. PubMed ID: 10219938 doi:10.1001/archinte.159.8.882

Mokkink, L.B., Terwee, C.B., Knol, D.L., Stratford, P.W., Alonso, J., Patrick, D.L., … de Vet, H.C. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology, 10*(1), 22. PubMed ID: 20298572 doi:10.1186/1471-2288-10-22

Mokkink, L.B., Terwee, C.B., Patrick, D.L., Alonso, J., Stratford, P.W., Knol, D.L., … de Vet, H.C.W. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology, 63*(7), 737–745. PubMed ID: 20494804 doi:10.1016/j.jclinepi.2010.02.006

Morris, J.N., Heady, J.A., Raffle, P.A., Roberts, C.G., & Parks, J.W. (1953). Coronary heart-disease and physical activity of work. *The Lancet, 262*(6796), 1111–1120. doi:10.1016/S0140-6736(53)91495-0

Nelson, M.E., Rejeski, W.J., Blair, S.N., Duncan, P.W., Judge, J.O., King, A.C., … Castaneda-Sceppa, C. (2007). Physical activity and public health in older adults: Recommendation from the American College of Sports Medicine and the American Heart Association. *Medicine & Science in Sports & Exercise, 39*(8), 1435–1445. PubMed ID: 17762378 doi:10.1249/mss.0b013e3180616aa2

Pedišić, Ž., & Bauman, A. (2015). Accelerometer-based measures in physical activity surveillance: Current practices and issues. *British Journal of Sports Medicine, 49*(4), 219–223. PubMed ID: 25370153 doi:10.1136/bjsports-2013-093407

Portney, L.G., & Watkins, M.P. (2014). *Foundations of clinical research: Applications to practice* (3rd ed.). Harlow, England: Pearson.

Prince, S.A., Adamo, K.B., Hamel, M.E., Hardt, J., Gorber, S.C., & Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity, 5*(1), 56. PubMed ID: 18990237 doi:10.1186/1479-5868-5-56

Prince, S.A., Cardilli, L., Reed, J.L., Saunders, T.J., Kite, C., Douillette, K., … Buckley, J.P. (2020). A comparison of self-reported and device measured sedentary behaviour in adults: A systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity, 17*(1), 31. PubMed ID: 32131845 doi:10.1186/s12966-020-00938-3

Rubio Castañeda, F.J., Tomás Aznar, C., & Muro Baquero, C. (2017). Validity, reliability and associated factors of the International Physical Activity Questionnaire adapted to elderly (IPAQ-E). *Revista Espanola De Salud Publica, 91,* e201701004.

Sagelv, E.H., Hopstock, L.A., Johansson, J., Hansen, B.H., Brage, S., Horsch, A., … Morseth, B. (2020). Criterion validity of two physical activity and one sedentary time questionnaire against accelerometry in a large cohort of adults and older adults. *BMJ Open Sport & Exercise Medicine, 6*(1), e000661. PubMed ID: 32153981 doi:10.1136/bmjsem-2019-000661

Sallis, J.F., & Saelens, B.E. (2000). Assessment of physical activity by self-report: Status, limitations, and future directions. *Research Quarterly for Exercise and Sport, 71*(Suppl.), 1–14. doi:10.1080/02701367.2000.11082780

Silsbury, Z., Goldsmith, R., & Rushton, A. (2015). Systematic review of the measurement properties of self-report physical activity questionnaires in healthy adult populations. *BMJ Open, 5*(9), e008430. PubMed ID: 26373402 doi:10.1136/bmjopen-2015-008430

Sylvia, L.G., Bernstein, E.E., Hubbard, J.L., Keating, L., & Anderson, E.J. (2014). A practical guide to measuring physical activity. *Journal of the Academy of Nutrition and Dietetics, 114*(2), 199–208. PubMed ID: 24290836 doi:10.1016/j.jand.2013.09.018

Szanton, S.L., Walker, R.K., Roberts, L., Thorpe, R.J., Wolff, J., Agree, E., … Seplaki, C. (2015). Older adults' favorite activities are resoundingly active: Findings from the NHATS study. *Geriatric Nursing, 36*(2), 131–135. PubMed ID: 25619566 doi:10.1016/j.gerinurse.2014.12.008

Tedesco, S., Barton, J., & O'Flynn, B. (2017). A review of activity trackers for senior citizens: Research perspectives, commercial landscape and the role of the insurance industry. *Sensors, 17*(6), 1277. doi:10.3390/s17061277

Tudor-Locke, C., Schuna, J.M., Barreira, T.V., Mire, E.F., Broyles, S.T., Katzmarzyk, P.T., & Johnson, W.D. (2013). Normative steps/day values for older adults: NHANES 2005–2006. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences, 68*(11), 1426–1432. PubMed ID: 23913932 doi:10.1093/gerona/glt116

Wen, C.P., Wai, J.P.M., Tsai, M.K., Yang, Y.C., Cheng, T.Y.D., Lee, M.-C., … Wu, X. (2011). Minimum amount of physical activity for reduced mortality and extended life expectancy: A prospective cohort study. *The Lancet, 378*(9798), 1244–1253. doi:10.1016/S0140-6736(11)60749-6

Westerterp, K.R. (2009). Assessment of physical activity: A critical appraisal. *European Journal of Applied Physiology, 105*(6), 823–828. PubMed ID: 19205725 doi:10.1007/s00421-009-1000-2

World Health Organization. (2010). *Global recommendations on physical activity for health*. Retrieved from https://www.who.int/dietphysicalactivity/publications/9789241599979/en/

World Health Organization. (2011). *Physical activity and older adults*. Retrieved from https://web.archive.org/web/20201127015524/https://www.who.int/teams/health-promotion/physical-activity/physical-activity-and-older-adults

APPENDIX 5: PAPER E (STUDY IV)

Larsen *et al. BMC Geriatrics* (2020) 20:412
https://doi.org/10.1186/s12877-020-01815-1

BMC Geriatrics

STUDY PROTOCOL Open Access

The MIPAM trial: a 12-week intervention with motivational interviewing and physical activity monitoring to enhance the daily amount of physical activity in community-dwelling older adults – a study protocol for a randomized controlled trial

Rasmus Tolstrup Larsen[1*], Christoffer Bruun Korfitsen[1,2], Carsten Bogh Juhl[3,4], Henning Boje Andersen[5], Jan Christensen[6,7] and Henning Langberg[8]

Abstract

Background: Physical Activity Monitors (PAMs) have been shown to effectively enhance level of physical activity (PA) in older adults. Motivational interviewing is a person-centred model where participants are guided using self-reflection and counselling, and addresses the behavioural and psychological aspects of why people initiate health behaviour change by prompting increases in motivation and self-efficacy. The addition of motivational interviewing to PA interventions may increase the effectiveness of PAMs for older adults.

Methods: This motivational interviewing and PA monitoring trial is designed as an investigator-blinded, two arm parallel group, randomized controlled superiority trial with primary endpoint after 12 weeks of intervention. The intervention group will receive a PAM-based intervention and motivational interviewing and the control group will only receive the PAM-based intervention. The primary outcome is PA, objectively measured as the average daily number of steps throughout the intervention period. Secondary outcome measures include self-reported PA health-related quality of life, loneliness, self-efficacy for exercise, outcome expectancy for exercise, and social relations. The outcomes will be analysed with a linear regression model investigating between-group differences, adjusted for baseline scores. Following the intention to treat principle, multiple imputation will be performed to handle missing values.

(Continued on next page)

* Correspondence: rala@sund.ku.dk
[1]Department of Public Health, Section of Social Medicine, University of Copenhagen, Copenhagen, Denmark
Full list of author information is available at the end of the article

(Continued from previous page)

Discussion: A moderate effect of daily PA measured using PAMs is expected in this superiority RCT investigating the effect of adding motivational interviewing to a PAM intervention. According to the World Health Organization, walking and cycling are key activities in regular PA and should be promoted. To increase the general public health and lower the burden of inactivity in older adults, cost-beneficial solutions should be investigated further. If this RCT shows that motivational interviewing can enhance the effect of PAM-based interventions, it might be included as an add-on intervention when appropriate. No matter what the results of this study will be, the conclusions will be relevant for clinicians as the dependence on technology is increasing, especially in relation to public health promotion.

Trial registration: NCT03906162, April 1, 2019.

Background

Twenty-seven percent of older adults in Denmark (65–74 years) and 39–46% of very old adults (age above 75) do not meet the World Health Organization's (WHO) recommendations for minimum physical activity (PA) [1] and the motivation for increasing PA is low for both age groups [2]. Physical inactivity and low PA levels have a major impact on global public health [3]. Physical inactivity among older adults is associated with disability and premature death and is one of the main barriers to healthy aging [4, 5]. Increased PA levels among older adults, including the ones living with chronic diseases, are associated with longevity benefits and healthy aging no matter the previous level of PA [5, 6].

Overall, PA promoting interventions do seem wo work well among older adults [7, 8] and furthermore, a review of reviews by Olanrewaju et al. found that behavioural and cognitive interventions are effective for increasing short-term PA in older adults [9]. Walking is the preferred form of PA among community-dwelling older adults [10], and participation in walking programs is an effective [9] means of increasing PA levels among this population. In order to maintain long-term participation in PA programs, individualized interventions based on theories of health behaviour change may be required [4, 9]. Social support may be important for increasing PA in older adults as social support and social networks influence health behaviours [11]. Lack of motivation for, or adherence to, exercise in older adults may be due to low self-efficacy or perceived barriers [12–15]. Physical activity monitors (PAMs) used to provide user feedback can facilitate motivational behavioural change and are often used in interventions to increase the average number of daily steps in older adults [16, 17].

However, PAMs might not be adequate or optimal for all older adults, as individualized goal-setting and social support have been reported as important factors in PA interventions [18]. A strategy including PA monitoring, goal setting [18] and Motivational Interviewing (MI) has been shown to promote maintenance of increased PA behaviour 6 months after intervention [19]. MI is a person-centred model where participants are guided using self-reflection and counselling [20]. MI addresses the behavioural and psychological aspects of why people initiate health behaviour changes by prompting increases in motivation and self-efficacy [21, 22]. In Denmark, MI is already well established among municipality health work with older adults or general practitioners' counselling of patients [23–29]. Furthermore, studies within older adults have reported MI to increase PA levels among patients with heart failure [30] and hip fracture [31]. Finally, older adults have reported the combination of PAMs and MI to be acceptable in a feasibility study aimed to keep people active after a fall management intervention, which to our knowledge is the only study that combines a PAM-based intervention with MI in older adults [32] Thus, MI shows potential for increasing PA levels and seems especially relevant to include and investigate in combination with a PAM-based intervention among Danish community-dwelling older adults [33, 34].

Objective

The objective of this RCT is to investigate the effect of MI as an add-on intervention to a PAM-based intervention measured by the average daily step count in community-dwelling older adults above the age of 70. It is hypothesized that: 1) MI will enhance the average daily step count among participants, 2) that MI will affect self-reported PA and quality of life, and 3) that self-efficacy and outcome expectancy for exercise will mediate this effect and explain heterogeneity in the results.

Trial design

The MIPAM trial is designed as an investigator-blinded, two arm parallel group, superiority RCT with primary endpoint after 12 weeks of intervention.

Methods

This protocol is reported according to the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT statement) [35].

Participants, interventions, and outcomes

Study setting

This RCT will be conducted nationwide among the community-dwelling older adults in Denmark.

Eligibility criteria

Participants will be considered eligible for inclusion if they: 1) are retired and community-dwelling, 2) are at least 70 years of age at the time of enrolment, 3) own a smartphone or tablet able to install the *Garmin Connect application,* 4) have an e-mail address and are able to correspond and complete the study survey, and 5) have hearing abilities sufficient to receive a telephone interview.

Participants will not be considered eligible for inclusion, and hence excluded, if: 1) they have cognitive impairment from moderate to severe dementia or Alzheimer's disease, 2) they are undergoing active chemotherapy or palliative care for cancer, and 3) or have a major mobility impairment preventing them from walking (e.g. from paralysis, amputations, severe arthrosis or arthritis, multiple sclerosis or Parkinson's disease).

Interventions

The control group will receive the PAM intervention and the experimental group will receive both the PAM intervention and a telephone-based MI intervention including goal setting for PA.

Physical activity monitor intervention

In a recent systematic review including 21 RCTs, PAMs has been shown to effectively enhance the daily number of daily steps in older adults [16, 17]. The PAM intervention consists of a PAM for everyday use in the intervention period and a pamphlet with information about Danish recommendations on PA in aging populations. The PAM will be the hip-worn Garmin Vivofit 3 monitor linked to a pre-specified *Garmin Connect* account. The participants will receive the PAMs and an installation guide, and will be asked to install the *Garmin Connect* application on their smartphone using a pre-specified ID/password in the app. The *Garmin Connect* application (https://connect.garmin.com/) allows participants to track, view and explore their daily step data. It allows for individual goal-setting on PA or other health related variables e.g. weight management, and it also allows the participants to connect with friends or relatives and create challenges with these. The participants will only be asked to install the application and use the automated goal-setting for daily steps, but they will be allowed to explore and use other functions of the application. Participants with installation difficulties will receive support by telephone from the research team. The participants will be asked to wear the monitor for all waking hours, except when bathing, every day for the 12-week intervention period.

Experimental intervention

The experimental intervention combines the PAM intervention with a MI-intervention, delivered by MI-trained physiotherapists (PT). During the 12-week intervention period, the participants will receive seven telephone calls. Using an intervention schedule inspired by the work of King et al. to facilitate initiation and maintenance of behaviour change, calls are delivered in the first, second, third, fifth, seventh, ninth and last intervention week [36].

In this person-centred intervention model, participants are guided through self-reflective counselling consistent with the MI approach [20]. They will receive feedback on their PA and health behaviours in relation to the national recommendations. Consistent with the original MI approach [20], this feedback will also highlight the discrepancy between their health goals and their current health behaviours.

The underlying theoretical perspective used to motivate the participants is derived from the Social Cognitive Theory (SCT) and the Transtheoretical Model (TTM) [37–40]. SCT proposes that to promote individuals' health behaviours, individuals must believe in their ability to carry out the specific behaviour, and they must also believe in its benefit [41, 42]. Self-efficacy and outcome expectations are key constructs and seen as direct predictors of PA behaviours, and they operate through indirect pathways affecting goal setting and the perception of sociostructural factors [37]. TTM was developed by Prochaska and DiClemente and posits that behaviour change follows a series of stages, which will be assessed by the counsellor; 1) precontemplation (individuals are not participating in any PA and have no intention to do so in the future), 2) contemplation (individuals are not participating in any PA but intend to start doing so in the next 6 months), 3) preparation (individuals intend to start participating in regular PA in the next 6 months and are starting to make small changes in their activity behaviour), 4) action (individuals meet defined criteria for PA but have done so for less than 6 months), and 5) maintenance (individuals have met defined criteria for PA for more than 6 months) [38–40]. A number of factors determine movement through the stages, including cognitive and behavioural processes of change, self-efficacy, and outcome expectancies.

Several theoretical constructs from the SCT and the TTM are addressed by the MI intervention. Personal factors and self-efficacy, in this setting for exercise, will be operationalized by coaching with realistic and measurable goal setting. Self-efficacy as a construct will be measured by the self-efficacy for exercise scale (SEE) [43]. Behavioural factors and outcome expectancies will be operationalized through discussion of benefits and barriers to health behavioural change, which should lead to increased perception of benefits and decreased perception of barriers [13]. Further, discussion of problem-solving approach to address behaviours will lead to an improved ability to do so. Outcome-expectancies will be measured by the Outcome-Expectancies for Exercise-2 scale (OEE-2) [44]. In the SCT, social support is an important construct for behavioural change. Environmental factors and social support will be operationalized through identification of supports for maintenance of health behavioural change, and specific goal setting for using supports, which will lead to increase level of support for the participant's health behavioural change. Participants will be encouraged to use a variety of supports including family and friends, as well as neighbourhood and community supports. In collaboration with local community partners, a community reference guide has been compiled that enables the counsellor to refer participants to specific community resources (e.g., walking groups).

Fidelity

The intervention (and the actual content of the motivational interviews) will be tailored to individual participants, but the number and timing of calls will not be adjusted. The project telephone counsellors are PTs with additional training and education in the MI approach to telephone health behaviour counselling. Training involved a four-day course, with reading materials, discussions with other study investigators and roleplaying exercises. During this study, with participants' verbal consent, telephone MI sessions will be audiotaped on a regular basis to ensure fidelity of intervention delivery and to provide counsellor feedback. Fidelity monitoring will be conducted by two coders that agree on a global score for each coded MI-session using the Motivational Interviewing Treatment Integrity Scale version 4 (MITI 4) [45]. The MITI 4 consists of four global ratings (Cultivating Change talk, Softening Sustain Talk, Partnership, and Empathy), and 10 individual behaviour counts (Questions, Simple Reflections, Complex Reflections, Persuade with Permission, Giving Information, Affirmations, Emphasize Autonomy, Seeking Collaboration, Persuade and Confront) which are counted in the time frame of the interview [45]. A median global score in each domain of four and a reflection to question ratio of more than one will be considered decent MI proficiency. Call completion, duration of the call, type of MI-intervention and stage of change will be noted after each telephone call (Fig. 1).

Outcomes

Primary outcome measure

The average number of steps per day during the first baseline week and the 12-week intervention period is the primary study outcome. The hip-worn Garmin Vivofit 3 tri-axial accelerometer will serve as the PAMs and thus measure the primary outcome. The commercially available Garmin Vivofit 3 has, to the best of our knowledge, only been validated in older adults by our own research group. The Garmin Vivofit 3 was validated with three other monitors and the hip-worn PAMs were found to be superior to wrist-worn PAMs among older adults with and without rollators [46].

Secondary outcome measures

Secondary outcome measures will include self-reported information from the participants on PA, health-related quality of life, loneliness, self-efficacy for exercise, outcome expectancy for exercise, and social relations. All secondary outcomes will be collected at baseline, at endpoint, and at six- and 12-month follow up.

The baseline self-report questionnaire will be completed before the intervention group receives the first motivational interview and the endpoint questionnaire will be distributed after 12 weeks of intervention and after the last motivational interview.

International physical activity questionnaire-short form (IPAQ-SF)

The seven-item IPAQ-SF assesses the amount of moderate to vigorous physical activity (MVPA), VPA, walking time and sedentary time, that has been performed in the past 7 days [47]. The score is categorized into three levels of PA; low, moderate and high [48]. A review of 16 international studies of the measurement properties of the IPAQ-SF assessment demonstrated acceptable reliability (Spearman's rho: 0,32-0,88) [47] and low to moderate concurrent validity compared to accelerometer with a pooled correlation coefficient of 0.30 (Spearman's rho range: 0,09-0,38) [49]. The Danish version has previously been used in a Danish population of older adults [50]. MVPA, walking time and sedentary time will be used as outcomes from the IPAQ-SF.

Nordic physical activity questionnaire short (NPAQ-short)

The two-item NPAQ-short [51] is a short revised version of the original NPAQ, a survey tool based on telephone interviews designed for the assessment of MVPA. It was developed to monitor compliance with the WHO recommendations on PA [52] and has showed moderate correlation with objectively measured MVPA (Spearman's rho: 0.33) in a Danish population with an average age of 43 (range: 17–85) [51]. Besides MVPA, the NPAQ-short produces four categories of PA according

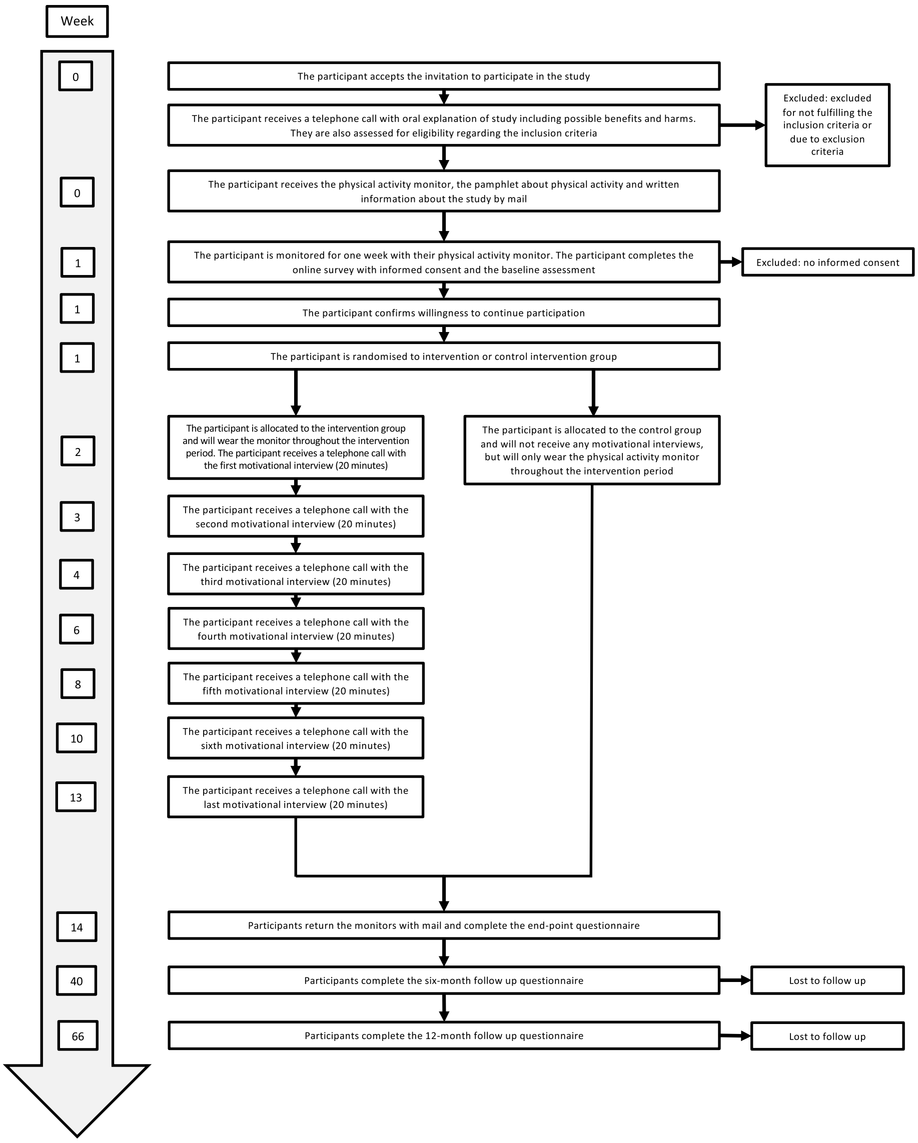


Fig. 1 SPIRIT participant timeline

to the WHO recommendations (inactive, insufficient physically active, sufficient physically active and optimally physically active). MVPA will be used as an outcome from the NPAQ-short.

The 5-level EuroQol-5 domain (EQ-5D-5L) quality of life questionnaire

The EQ-5D-5L is a generic Health-Related Quality of Life (HRQoL) measurement tool developed as a non-disease-specific instrument for HRQoL [53, 54]. EQ-5D-5L comprises of five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), each of which has five levels (no problems, slight problems, moderate problems, severe problems or unable to), and a visual analogue scale (EQ VAS) [54]. The EQ VAS records the patient's self-rated health on a vertical visual analogue scale, where the endpoints are labelled 'The best health you can imagine' and 'The worst health you can imagine'. The EQ-5D-5L has shown general feasibility for measuring HRQoL in a population sample of older adults [55]. The test-retest reliability have been evaluated for the EQ-5D-3L index (correlation: 0.67) and the EQ-5D-3L VAS (0.53) [56] but not for the EQ-5D-5L. The EQ-5D-3L has shown fair to moderate convergent validity by correlations with five related domains of the WHO-5 (Spearman's rho: 0.29–0.61) [57]. The EQ-5D-5L is adapted to Danish [58] but no psychometric evaluation of the Danish version has been published. The EQ VAS score will be used as an outcome from the EQ-5D questionnaire.

UCLA loneliness scale

The 20-item UCLA loneliness scale (third version) is a self-report measure of loneliness and social isolation [59]. The scale consists of 11 positive and nine negative items and the total score is calculated as the sum of 20 items (0–60), with a higher score indicating more loneliness. The negative items (one, five, six, nine, 10, 15, 16, 19 and 20) are reversed before the scores are summed (i.e. high score equals less loneliness). The scale is adapted to Danish (translation found in Additional file 2) and has shown high internal consistency (Cronbach's Alpha: 0.92) and moderate convergent validity with other measures of emotional loneliness (r: 0.69) and social loneliness (r: 0.73). In addition, the scale has showed moderate discriminant validity in relation to self-esteem (r: − 0.58), depression (r: 0,59), extraversion (r: 0.57) and neuroticism (r: 0,58). In a population of older adults, the scale has shown good internal consistency (Cronbach's Alpha: 0.87) [60]. The total score will be used as an outcome from the UCLA loneliness scale.

Self-efficacy for exercise

The nine-item SEE addresses confidence to engage in regular exercise [43], when challenged by known barriers to exercise [61]. The scale was initially developed for sedentary adults living in the community who participated in an outpatient exercise program [62] and was revised to be applicable to older adults [43]. Response categories range from 0 (no confidence) to 10 (very confident) [43]. Item scores are used to calculate a total score (0–90), with higher scores indicating higher confidence, or self-efficacy, related to exercise. The SEE-DK has been translated and adapted to Danish community-dwelling older adults by our research group (translation found in Additional file 4). The average score will be used as an outcome from the SEE-DK.

Outcome expectancy for Exercise-2

The 13-item OEE-2 scale was developed from the original 9-item Outcome Expectations for Exercise (OEE) scale that focused on measuring the positive outcome expectations for exercise (POEE). Based on qualitative findings [61, 63], the original OEE was revised to include four items that focused on negative outcome expectations for exercise (NOEE) [44]. It was initially developed for older adults [64, 65]. To complete the OEE2-DK scale the participants are asked, using a Likert scale, to *strongly agree*, *agree*, *neither agree nor disagree*, *disagree*, or *strongly disagree* with each statement of exercising. The POEE and NOEE subscales are scored by calculating the average score on each scale (1–5) and the items three, six, nine and 12 (NOEE subscale) are reversed before the scores are summed [44]. The OEE2-DK has been translated and adapted to Danish community-dwelling older adults by our research group (translation found in Additional file 3). The average score will be used as an outcome from the OEE2-DK.

These secondary outcomes be completed as will be completed as follow-up measures six and 12 months after ending the intervention. They will be conducted as online surveys.

Social- and demographic baseline variables

The 42-item Copenhagen Social Relations Questionnaire (CRSQ), will be used to describe participants in terms of structural and social relations. CRSQ was originally developed in Danish in 1999 [66] and measures the structural aspects of social relations, with a focus on frequency and diversity of social contact, and functional aspects with focus on perceived social support. CSRQ has been used in several Danish population-based surveys including in the Copenhagen Aging and Midlife Biobank (CAMB) [67]. In a sample of 38- to 69-year-old adults, the CSRQ showed acceptable face and content validity and good test-retest reliability, with 41% of the items achieving substantial to almost perfect agreement (kappa: 0.65–0.97) and the rest showing moderate agreement (kappa: 0.41–0.60) [68]. The CRSQ will be used to report if the participants are living alone.

Table 1, including socio-demographics of included participants, and Table 2, including PA characteristics, will be used to report relevant baseline information on the participants.

Sample size and power considerations

The estimated number of participants required is 128. This number will be sufficient to show a 0.5 standard deviation difference between groups, equal to a moderate effect size, on the primary outcome (steps per day). The number of participants will yield a power on 80% with a significance level of 0.05. To account for participation attrition, this study will enrol 20% more participants than required, for total of 154 participants divided into two comparison groups.

Recruitment

We will recruit participants through online advertisements on Facebook and LinkedIn, in non-profit organizations working with older adults (such as activity organizations) and at activity centres and other communities of older adults. Participants eligible for inclusion will receive the information necessary for participation by mail and complete online questionnaires. The participants will only have contact with the research team via phone or e-mail correspondence.

Table 1 Socio-demographics of included participants

Characteristics	Overall (n=)	Control group (n=)	Intervention group (n=)	*p*
Age in years, mean (95% CI)	–	–	–	⊗
Male, n (%)	–	–	–	⊠
BMI, mean (95%CI)	–	–	–	⊗
Education	–	–	–	⊠
No education, n (%)	–	–	–	
Primary or secondary education, n (%)	–	–	–	
Tertiary education, n (%)	–	–	–	
Master's degree, n (%)	–	–	–	
Living alone, n (%)	–	–	–	⊠
Long-term illness, injuries or disability more than 6 months, n (%)	–	–	–	⊠
Smoking	–	–	–	
Smokes**,** n (%)	–	–	–	⊠
Quit smoking**,** n (%)	–	–	–	
Never smoked, n (%)	–	–	–	
Wants to be more physically active	–	–	–	
Yes, n (%)	–	–	–	⊠
No, n (%)	–	–	–	
Do not know, n (%)	–	–	–	
Uses a PAM before enrolment, n (%)	–	–	–	⊠
Walking aids				
No walking aids, n (%)	–	–	–	⊠
Cane user, n (%)	–	–	–	
Rollator user, n (%)	–	–	–	
Reports to be in pain, n (%)	–	–	–	⊠
EQ-5D	–	–	–	
Mobility – reporting problems n (%)	–	–	–	⊠
Self-Care – reporting problems n (%)	–	–	–	⊠
Usual Activity – reporting problems n (%)	–	–	–	⊠
Pain/Discomfort – reporting problems n (%)	–	–	–	⊠
Anxiety/Depression – reporting problems n (%)	–	–	–	⊠
EQ VAS, mean (95%CI)	–	–	–	⊗
UCLA Loneliness Scale, mean n (%)	–	–	–	⊗
OEE average score, mean (95% CI)	**–**	**–**	**–**	⊗
SEE average score, mean (95% CI)	–	–	–	⊗

BMI: Body Mass Index, PAM: Physical Activity Monitor, EQ-5D: EuroQol Research Foundation Five Domains, UCLA: University of California Los Angeles, OEE: Outcome Expectancy for Exercise, SEE: Self Efficacy for Exercise. IPAQ-SF: International Physical Activity Questionnaire-Short Form, NPAQ: Nordic Physical Activity Questionnaire-Short Form, MVPA: Moderate to Vigorous Physical Activity, SD: standard deviation, 95% CI: 95% confidence intervals, IQR: interquartile range, ⊗ Test for between-group difference with unpaired t-test, ⊙ Test for between group difference with Mann-Whitney U test, ⊠ Test for between group difference with Chi^2 test, *p* values for between group difference ≤.05 are considered significant

Assignment of interventions
Allocation

Sequence generation and allocation concealment mechanism Participants will be randomly assigned to either the intervention or the control group, with a 1:1 allocation. After completion of the one-week baseline period, eligible participants will be randomized into blocks consisting of a minimum four participants, stratified on sex (M/F) and average daily baseline step count for the baseline period. Randomization of participants will be performed every week, except for weeks with less than four new participants.

Table 2 Physical activity characteristics of included participants

Characteristics	Overall (n=)	Control group (n=)	Intervention group (n=)	p
Baseline Physical Activity Average daily step count, mean (95% CI)	–	–	–	⊗
IPAQ-SF	–	–	–	
MVPA, mean (95% CI)	–	–	–	⊗
MPA, mean (95% CI)	–	–	–	⊗
VPA, mean (95% CI)	–	–	–	⊗
Walking time, mean (95% CI)	–	–	–	⊗
Sedentary time, mean (95% CI)	–	–	–	⊗
Low activity level, n (%)	–	–	–	
Moderate activity level, n (%)	–	–	–	⊠
High activity level, n (%)	–	–	–	
NPAQ	–	–	–	
MVPA, mean (95% CI)	–	–	–	⊗
VPA, mean (95% CI)	–	–	–	⊗
Physically inactive, n (%)	–	–	–	
Insufficient physically active, n (%)	–	–	–	
Sufficient physically active, n (%)	–	–	–	⊠
Optimally physically active, n (%)	–	–	–	

IPAQ-SF: International Physical Activity Questionnaire-Short Form, NPAQ: Nordic Physical Activity Questionnaire-Short Form, MVPA: Moderate to Vigorous Physical Activity, SD: standard deviation, 95% CI: 95% confidence intervals, IQR: interquartile range, ⊗ Test for between-group difference with unpaired t-test, ⊙ Test for between group difference with Mann-Whitney U test, ⊠ Test for between group difference with Chi^2 test, p values for between group difference ≤.05 are considered significant

Participants will be randomized using the statistical software package STATA. Allocation concealment will be ensured, as the allocation will not be available until the patient has been recruited into the trial, which takes place after the baseline step count measurements have been completed.

One investigator will oversee the randomization. That investigator will receive a list of participant IDs every week and randomize the participants according to the above method. This investigator will not have any role in recruitment or in statistical analyses. The data-analysis-responsible investigator will be blinded for participant allocation. As the primary outcome is objectively measured steps per day, the outcome assessor of the primary outcome can be considered blinded. As the secondary outcome measures are self-reported, the outcome assessor is not blinded.

Due to the nature of the intervention, both participants and staff conducting the motivational interview in the intervention group will not be blinded to allocation. However, they will be strongly encouraged to not disclose the allocation status of the participants with the principal investigator who will conduct the analyses. The group names of the intervention and the control group will be anonymized before the data will be analysed to ensure blinding of the principal investigator.

Data collection, management, and analysis

Data collection methods

This section includes plans for assessment and collection of outcomes.

Primary outcome The primary outcome (average steps per day in the 12-week intervention period) will be extracted from the data management software program Fitrockr. Participants will be asked to synchronize their PAMs and their Garmin Connect application daily, ensuring daily storage of the step counts. Every week, participants who fail to synchronize their PAM will be reminded via e-mail or by telephone. The PAMs have the capacity to store the step counts for 30 days; therefore, no data loss is anticipated, even if the participants fail to synchronize their PAM for longer periods of time. Fitrockr will extract the data from Garmin Connect and make daily step counts available for export through their service. When the participant has completed the 12-week intervention, the daily totals will be extracted as 84 variables (12*7). After the data extraction, the average daily step count will be calculated.

Secondary outcomes All secondary outcomes are participant-reported and administered through the online survey platform SurveyXact. All participants will receive an email with an electronic SurveyXact invitation on the day of randomization. On the last day of intervention (day 84), the participants will receive a similar SurveyXact invitation with the end-point questionnaire. The six- and 12-month follow up assessments will be administered in similar ways as the end-point assessment.

Demographic and other baseline items Non-outcome variables will be included in the baseline questionnaire and thus be self-reported. These variables include: sex (male/female), age in years, height in cm, weight in kg, highest completed education (no education, primary education, secondary education, tertiary education), marital status (married, widow/widower, single, divorced), smoking habits (never smoked, former smokers, smoker), present pain (yes/no), long-term illness or disability from injury (more than 6 months yes/no), felt limited in performing daily activities because of health issues (seriously limited, somewhat limited, not limited), use of walking aids (no walking aids, cane or rollator), use of PAMs before enrolment (yes/no), would like to be more physically active (yes, no, do not know).

Reasons for dropout will be collected from each discontinued participant by the primary investigator after discontinuation.

Data management

All outcomes data will be collected and stored electronically. No personal data will be exported from Fitrockr or SurveyXact without pseudonymization. Complete anonymization of all data will be performed after the last follow-up period. Data protection agency approval Reference number: 514–0268/18–3000.

Steps per day will be stored each time the participant synchronizes the PAM. The data-handling-responsible program Fitrockr will extract data from the Garmin applications and store these data according to the agreements. When a participant completes the intervention period, their data will be exported from the Fitrockr database and stored securely at the University of Copenhagen server.

Statistical methods

Distributions of continuous data will be evaluated by inspecting Quantile-Quantile plots of the standardized residuals and histograms with normal distribution curves. Continuous data with normal distributions will be analysed with parametric statistics. Continuous data with non-normal distributions will be analysed as ordinal data with non-parametrical statistics. Categorical data will be presented as frequencies.

The primary outcome, average daily step count, will be analysed with a linear regression model investigating the between-group differences, adjusted for sex (M/F) and baseline daily step count. Following the intention to treat principle, the Gaussian normal regression method will be used to impute missing values (multiple imputation on baseline step count, gender and age).

The same procedure will be used to analyse between group differences on secondary outcomes and all secondary outcomes will be adjusted for baseline score, baseline daily step count and sex (M/F). Harms will be evaluated by calculating the relative risk (RR), separately for serious and non-serious adverse event between the intervention and control group.

In calculating the average daily step count, days with less than 100 steps will be handled as “days of non-wear” and excluded assessing the mean step count.

Table 3 is the outline table for the reporting of end-point values for primary and secondary outcomes.

Table 3 End-point values for primary and secondary outcomes, adjusted for sex, baseline scores and baseline daily step count

Characteristics	Control group (n=)	Intervention group (n=)	*P*
Primary outcome Average daily step count, mean (95% CI)	–	–	–
Secondary outcomes	–	–	–
EQ VAS, mean (95% CI)	–	–	–
UCLA Loneliness Scale, mean n (%)	–	–	–
OEE average score, mean (95% CI)	–	–	–
SEE average score, mean (95% CI)	–	–	–
IPAQ-SF MVPA, mean (95% CI)	–	–	–
IPAQ-SF Walking time, mean (95% CI)	–	–	–
IPAQ-SF Sedentary time, mean (95% CI)	–	–	–
NPAQ MVPA, mean (95% CI)	–	–	–

BMI: Body Mass Index, PAM: Physical Activity Monitor, EQ-5D: EuroQol Research Foundation Five Domains, UCLA: University of California Los Angeles, OEE: Outcome Expectancy for Exercise, SEE: Self Efficacy for Exercise. IPAQ-SF: International Physical Activity Questionnaire-Short Form, NPAQ: Nordic Physical Activity Questionnaire-Short Form, MVPA: Moderate to Vigorous Physical Activity, SD: standard deviation, 95% CI: 95% confidence intervals, IQR: interquartile range, between-group differences calculated from linear regression model, adjusted for baseline scores, sex and baseline daily step count, using imputed values from the Gaussian normal regression method (baseline step count, sex and age). p values for between group difference $\leq .05$ are considered significant

Monitoring

Data monitoring

Not applicable/relevant.

Harms

In our study, adverse events will be defined as any unintended negative consequences in a participant without regard to the possibility of a causal relationship with the intervention. Adverse event rates will be measured after the subject has provided consent and enrolled in the study. All adverse events occurring after entry into the study will be recorded. The participants will be asked at the end-point questionnaire if they experienced any adverse events in terms of using the PAMs or trying to enhance their daily amount of PA.

Auditing

No auditing has been protocolled.

Ethics and dissemination

Research ethics approval

According to written correspondence with the Danish Ethics Committee in the Capital Region of Denmark, this trial was not subject to the current laws on research ethics in Denmark due to the non-invasive behavioural change intervention. Thus, this study was pre-approved and can be conducted without further approval from the Danish Ethics Committee of the Capital Region of Denmark (Journal-nr.:18004960).

Protocol amendments

Any modifications to the protocol which may affect the study procedure, potential benefit to participants, or may affect safety, including changes of study objectives, study design, sample population, sample size, study procedures, or significant administrative aspects will require a formal amendment to the protocol that will be revised and re-uploaded to Clinicaltrials.gov.

Consent or assent

Informed consent will be collected electronically via SurveyXact. Prior to agreeing and signing the consent survey, the participant will receive written information about the study by email. If the participant has any questions they may contact the study-responsible researcher. The participant is informed orally and in writing that they can withdraw their consent at any time without affecting current or future treatment in the Danish healthcare system. The translated version of the informed consent material can be found in Additional file 1.

Confidentiality

All study-related information and collected data on participants will be stored securely on a server at University of Copenhagen. All extractions from this server will be followed by immediately anonymization of the dataset.

Discussion

Implications

We expect a clinically relevant moderate effect on PA from the experimental intervention in this RCT. According to the World Health Organization, walking and cycling are key activities in regular PA and should be promoted among older adults [69–71]. To increase the general public health and lower the burden of inactivity in older adults, the efficacy of cost-beneficial solutions should be investigated further [70]. If this RCT shows that MI can enhance the effect of PAM-based interventions, it might be included as a cost-benefit add-on intervention when appropriate. The conclusions from this study will be relevant for clinicians as the dependence on technology is increasing, especially in relation to public health promotion.

Methodology

Several recommendations for conducting clinical trials have been published [72, 73] and following the SPIRIT [35] reporting framework does not mean that the trial will be effective. In this section, the most relevant pitfalls in conducting this particular RCT will be discussed.

Unclear hypotheses and multiple objectives often hinder clinical trials as they may confuse readers and lower the applicability [73]. To answer the specific research question about adding MI to PAM-based interventions, this RCT uses a simple design to increase the generalizability of findings outside of the trial context.

If clinically irrelevant outcomes are used, the trail may not reflect the real world concerns of clinicians, which may affect applicability of the trial [74]. Often, surrogate outcomes are used to show an effect if the intervention fails to change the real and clinically relevant outcomes [74]. Thus, HRQoL, outcome-expectancy and self-efficacy of the participants may improve among the intervention group participants and might be considered as positive changes. However, because of the primary outcome of interest, daily PA, is a distinct construct, the secondary outcomes will only be used to explain the effect (or heterogeneity of) in the primary outcome. Daily steps have been shown to be a critical construct as it is highly associated with longevity and health status among older adults [75, 76]. In this study, daily steps will be validly measured by the Garmin Vivofit 3 [46], and thus able to serve as a critical outcome relevant to both clinicians and decision makers.

When selecting eligibility criteria for study participation, researchers should consider whether to strive for a homogeneous or heterogenous sample. For study enrolment purposes, achieving a homogenous sample is more challenging, but may also result in a more precise effect estimate. A heterogenous sample is expected to increase the generalizability of results in exchange for less precise effect estimates and reduced ability to draw conclusions from results [73]. In this RCT, we have chosen eligibility criteria that allow for a quite heterogenic group of study participants, as we expect our results to be most affected by self-efficacy for exercise and outcome expectancy, rather than participant demographics factors [12, 77]. Our study participants may react differently to the behavioural change intervention, but this will most likely reflect the real-life situations where a single approach may not apply equally [4]. In summary, we have chosen generalizability of the results in favour of effect estimate precision.

When choosing the comparator, the control group intervention content, should be relevant and active. Both the PREPARE guideline, CONSORT statement and the SPIRIT checklist suggest building the intervention and control intervention content on a recent published systematic review [35, 72, 78]. Our study group recently published a systematic review that found that PAM-based interventions are effective and that future comparison studies should not use passive control groups to investigate the effect any further [16]. Instead, future studies should investigate the right and relevant questions, such as "does PAM-based interventions work better than …? " or "can we enhance the effect of a PAM-based intervention by? " [16]. This RCT stands as a superiority trial investigating if MI should be added to PAM-based interventions among older adults.

Selecting an appropriate study timeline to measure an effect, and in this trial, long-term behavioural change, is critical in trial design [72]. We considered both the practical possibilities and the optimal intervention length and arrived on a 12-week intervention period. If the intervention, and thus exposure to MI is too short the intervention is unlikely to demonstrate positive outcomes. However, a prolonged intervention may hinder implementation in a real-world setting. Among the 21 PAM-based intervention studies included in the previously mentioned systematic review, the median intervention time was also 12 weeks (range 4 to 52). More importantly, the intervention length was not associated with effect size [16]. However, researchers must also include follow-up time to ensure long-term adherence to the health behaviour change, which is ultimately the desired outcome of MI [79]. This RCT will use self-reported measures to investigate long-term adherence to the behaviour change after six- and 12-months after the intervention. Results from these long-term follow up periods will also be published and are expected to be relevant. This is because it is hypothesized that participants who received MI will develop more effective strategies to ensure long-term adherence, compared to the participants who only received the PAM-based intervention.

Conclusions

PAMs has been shown to effectively enhance PA-levels among older adults and passive comparisons are therefore not encouraged. Future research should investigate whether the effect of PAMs can be enhanced by adding relevant behavioural change content, such as MI, in populations of older adults. This RCT will be conducted according to current best practice guidelines and will help future clinicians and decision makers to decide if MI should be included in PAM-based interventions among older adults.

Declaration of interests

The content presented within this protocol and the study was produced as part of the project REACH: this project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 690425.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12877-020-01815-1.

Additional file 1. Informed consent materials.
Additional file 2. Danish translation of UCLA.
Additional file 3. Danish translation of OEE-2.
Additional file 4. Danish translation of SEE.

Abbreviations

PAMs: Physical Activity Monitors; PA: Physical Activity; WHO: World Health Organization; MI: Motivational Interviewing; SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials; PT: Physiotherapist; SCT: Social Cognitive Theory; TTM: Transtheoretical Model; SEE: Self-efficacy for Exercise scale; OEE-2: Outcome-Expectancies for Exercise-2 scale; MVPA: Moderate to Vigorous Physical Activity; IPAQ-SF: International Physical Activity Questionnaire-Short Form; NPAQ-short: Nordic Physical Activity Questionnaire short; EQ-5D-5L: 5-level EuroQol-5 Domain; HRQoL: Health-Related Quality of Life; EQ VAS: Visual Analogue Scale; CRSQ: Copenhagen Social Relations Questionnaire; RR: Relative Risk

Acknowledgements

The authors would like to thank Dr. Luke Turcotte, for carefully reading and editing the manuscript in its final version.

Authors' contributions

All authors did actively contribute to the design of this intervention study. The intervention content was developed by RTL and CBK. The outcome measures were chosen and evaluated by RTL, CBK, CJ, JC and HL. Besides being actively contributing to the design, HBA served as the primary technology expert and CJ served as the methodological expert as well as proving statistical counselling. The study was originally initiated by HL and HBA. RTL was the primary author of this protocol, but all authors contributed to its content. All authors have read and approved the manuscript. *Fitrockr (https://www.fitrockr.com/#!/welcome)*, will serve as the data management team responsible for handling the data export from the PAMs and exporting the data to the research and intervention team. Data processor agreement forms have been completed prior to registration and beginning of this RCT, in accordance with approval from the Danish Data Protection Agency.

Funding
The content presented within this paper was produced as part of the project REACH: this project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 690425 [80]. The funder (Horizon 2020) have not contributed to any work regarding this study protocol.

Availability of data and materials
The datasets that will be used and analysed during the study will be available from the corresponding author on reasonable request.

Ethics approval and consent to participate
According to written correspondence with the Danish Ethics Committee in the Capital Region of Denmark, this trial was not subject to the current laws on research ethics in Denmark due to the non-invasive behavioural change intervention. The study can be conducted without further approval from the Danish Ethics Committee of the Capital Region of Denmark (Journal-nr.:18004960).

Consent for publication
Not applicable.

Competing interests
All authors declare to have no financial or non-financial competing interests.

Author details
[1]Department of Public Health, Section of Social Medicine, University of Copenhagen, Copenhagen, Denmark. [2]Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark. [3]Research Unit of Musculoskeletal Function and Physiotherapy, Institute of Sports Science and Clinical Biomechanics, Faculty of Health Sciences, University of Southern Denmark, Odense, Denmark. [4]Department of Physiotherapy and Occupational Therapy, Copenhagen University Hospital, Gentofte, Denmark. [5]Technical University of Denmark, DTU Management Engineering Inst, Lyngby, Denmark. [6]Department of Occupational- and Physiotherapy, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark. [7]National Centre for Rehabilitation and Palliative Care, University of Southern Denmark and Odense University Hospital, Odense, Denmark. [8]Department of Public Health, Section for Health Services Research, University of Copenhagen, Copenhagen, Denmark.

Received: 12 September 2019 Accepted: 5 October 2020
Published online: 20 October 2020

References

1. WHO | The Global Strategy on Diet, Physical Activity and Health (DPAS). WHO. http://www.who.int/nmh/wha/59/dpas/en/. Accessed 17 Jan 2019.
2. Danskernes Sundhed - Den Nationale Sundhedsprofil 2017. https://www.sst.dk/da/udgivelser/2018/danskernes-sundhed-den-nationale-sundhedsprofil-2017. Accessed 17 Jan 2019.
3. Lee I-M, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Impact of physical inactivity on the World's major non-communicable diseases. Lancet. 2012;380:219–29.
4. Bangsbo J, Blackwell J, Boraxbekk C-J, Caserotti P, Dela F, Evans AB, et al. Copenhagen consensus statement 2019: physical activity and ageing. Br J Sports Med. 2019;53:856–8.
5. Daskalopoulou C, Stubbs B, Kralj C, Koukounari A, Prince M, Prina AM. Physical activity and healthy ageing: a systematic review and meta-analysis of longitudinal cohort studies. Ageing Res Rev. 2017;38:6–17.
6. Mok A, Khaw K-T, Luben R, Wareham N, Brage S. Physical activity trajectories and mortality: population based cohort study. BMJ. 2019;365:l2323.
7. Higgins JPT, Green S. 12.6.2 re-expressing SMDs using rules of thumb for effect sizes. In: Cochrane Handbook for Systematic Reviews of Interventions; 2011. https://handbook-5-1.cochrane.org/chapter_12/12_6_2_re_expressing_smds_using_rules_of_thumb_for_effect_sizes.htm. Accessed 20 May 2020.
8. Grande GD, Oliveira CB, Morelhão PK, Sherrington C, Tiedemann A, Pinto RZ, et al. Interventions promoting physical activity among older adults: a systematic review and meta-analysis. The Gerontologist. 2019. https://doi.org/10.1093/geront/gnz167.
9. Olanrewaju O, Kelly S, Cowan A, Brayne C, Lafortune L. Physical activity in community dwelling older people: a systematic review of reviews of interventions and context. PLoS One. 2016;11. https://doi.org/10.1371/journal.pone.0168614.
10. Szanton SL, Walker RK, Roberts L, Thorpe RJ, Wolff J, Agree E, et al. Older adults' favorite activities are resoundingly active: findings from the NHATS study. Geriatr Nur (Lond). 2015;36:131–5.
11. Thomas M, Victor CR, Sullivan MP. Understanding and alleviating loneliness in later life: perspectives of older people. Qual Ageing Older Adults. 2016;17:168–78.
12. Shieh C, Weaver MT, Hanna KM, Newsome K, Mogos M. Association of Self-Efficacy and Self-Regulation with nutrition and exercise behaviors in a community sample of adults. J Community Health Nurs. 2015;32:199–211.
13. Williams DM, Anderson ES, Winett RA. A review of the outcome expectancy construct in physical activity research. Ann Behav Med Publ Soc Behav Med. 2005;29:70–9.
14. Cohen-Mansfield J, Marx MS, Guralnik JM. Comparison of exercise models in an elderly population. Aging Clin Exp Res. 2006;18:312–9.
15. Ziebart C, McArthur C, Lee L, Papaioannou A, Laprade J, Cheung AM, et al. "Left to my own devices, I don't know": using theory and patient-reported barriers to move from physical activity recommendations to practice. Osteoporos Int. 2018;29:1081–91.
16. Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance amount of physical activity in older adults – a systematic review and meta-analysis. BMC Eur J Aging Phys Act. 2019; In press.
17. Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance the daily amount of physical activity in elderly—a protocol for a systematic review and meta-analysis. Syst Rev. 2018;7:69.
18. Victor CR, Rogers A, Woodcock A, Beighton C, Cook DG, Kerry SM, et al. What factors support older people to increase their physical activity levels? An exploratory analysis of the experiences of PACE-lift trial participants. Arch Gerontol Geriatr. 2016;67:1–6.
19. Eakin E, Reeves M, Winkler E, Lawler S, Owen N. Maintenance of physical activity and dietary change following a telephone-delivered intervention. Health Psychol Off J Div Health Psychol Am Psychol Assoc. 2010;29:566–73.
20. Miller WR, Rollnick S. Motivational interviewing: helping people change. 3rd ed. New York: Guilford Press; 2013.
21. Cummings SM, Cooper RL, Cassie KM. Motivational interviewing to affect behavioral change in older adults. Res Soc Work Pract. 2009;19:195–204.
22. Hardcastle SJ, Hancox J, Hattar A, Maxwell-Smith C, Thøgersen-Ntoumani C, Hagger MS. Motivating the unmotivated: how can health behavior be changed in those unwilling to change? Front Psychol. 2015;6. https://doi.org/10.3389/fpsyg.2015.00835.
23. Municipality of Odense. Odense Kommune sender 3500 medarbejdere på uddannelse i... The Municipality of Odense educates 3500 employees in Motivational Interviewing. 2015. https://www.odense.dk/presse/pressemeddelelser/pressemeddelelser-2014/odense-kommune-sender-3500-medarbejdere-paa-uddannelse-i-motiverende-bogersamtale. Accessed 23 Jun 2020.
24. Rodnick JF. Denmark: teaching GPs to use motivational interviewing. Fam Med. 2007;39:215.
25. Rubak S, Sandbæk A, Lauritzen T, Christensen B. Motivational interviewing: a systematic review and meta-analysis. Br J Gen Pract. 2005;55:305.
26. Lindhardt CL, Rubak S, Mogensen O, Hansen HP, Goldstein H, Lamont RF, et al. Healthcare professionals experience with motivational interviewing in their encounter with obese pregnant women. Midwifery. 2015;31:678–84.
27. Charles M, Bruun NH, Simmons R, Dalsgaard E-M, Witte D, Jorgensen M, et al. The effect of training GPs in motivational interviewing on incident cardiovascular disease and mortality in people with screen-detected diabetes. Results from the ADDITION-Denmark randomised trial. BJGP Open. 2020;4:1–3.
28. Rosenbek Minet LK, Wagner L, Lønvig EM, Hjelmborg J, Henriksen JE. The effect of motivational interviewing on glycaemic control and perceived competence of diabetes self-management in patients with type 1 and type 2 diabetes mellitus after attending a group education programme: a randomised controlled trial. Diabetologia. 2011;54:1620–9.
29. Brandt CJ, Søgaard GI, Clemensen J, Sndergaard J, Nielsen JB. General practitioners' perspective on eHealth and lifestyle change: qualitative interview study. JMIR MHealth UHealth. 2018;6:e88.
30. Brodie DA, Inoue A. Motivational interviewing to promote physical activity for people with chronic heart failure. J Adv Nurs. 2005;50:518–27.

31. O'Halloran PD, Shields N, Blackstock F, Wintle E, Taylor NF. Motivational interviewing increases physical activity and self-efficacy in people living in the community after hip fracture: a randomized controlled trial. Clin Rehabil. 2016;30:1108–19.
32. Audsley S, Kendrick D, Logan P, Jones M, Orton E. A randomised feasibility study assessing an intervention to keep adults physically active after falls management exercise programmes end. Pilot Feasibility Stud. 2020;6:37.
33. Lundahl BW, Kunz C, Brownell C, Tollefson D, Burke BL. A meta-analysis of motivational interviewing: twenty-five years of empirical studies. Res Soc Work Pract. 2010;20:137–60.
34. Lundahl B, Moleni T, Burke BL, Butters R, Tollefson D, Butler C, et al. Motivational interviewing in medical care settings: a systematic review and meta-analysis of randomized controlled trials. Patient Educ Couns. 2013;93:157–68.
35. Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 statement: defining standard protocol items for clinical trials. Ann Intern Med. 2013;158:200–7.
36. King AC, Haskell WL, Taylor CB, Kraemer HC, DeBusk RF. Group- vs home-based exercise training in healthy older men and women: a community-based clinical trial. JAMA. 1991;266:1535–42.
37. Luszczynska A, Schwarzer R. Social cognitive theory. Fac Health Sci Publ. 2015:225–51.
38. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: applications to addictive behaviors. Am Psychol. 1992;47:1102–14.
39. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol. 1983;51:390–5.
40. Marshall SJ, Biddle SJ. The transtheoretical model of behavior change: a meta-analysis of applications to physical activity and exercise. Ann Behav Med Publ Soc Behav Med. 2001;23:229–46.
41. Bandura A. Self-efficacy: the exercise of control. New York: W H Freeman/ Times Books/ Henry Holt & Co; 1997.
42. Bandura A. Self-Efficacy in Changing Societies. Self-Efficacy in Changing Societies. New York: Cambridge University Press; 1995.
43. Resnick B, Jenkins LS. Testing the reliability and validity of the self-efficacy for exercise scale. Nurs Res. 2000;49(3):154–9. https://journals.lww.com/nursingresearchonline/Fulltext/2000/05000/Testing_the_Reliability_and_Validity_of_the.7.aspx.
44. Resnick B. Reliability and validity of the outcome expectations for exercise Scale-2. J Aging Phys Act. 2005;13:382–94.
45. Moyers TB, Rowell LN, Manuel JK, Ernst D, Houck JM. The motivational interviewing treatment integrity code (MITI 4): rationale, preliminary reliability and validity. J Subst Abus Treat. 2016;65:36–42.
46. Larsen RT, Korfitsen CB, Juhl CB, Andersen HB, Langberg H, Christensen J. Criterion validity for step counting in four consumer-grade physical activity monitors among older adults with and without rollators. Eur Rev Aging Phys Act. 2020;17:1.
47. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc. 2003;35:1381–95.
48. IPAQ scoring protocol - International Physical Activity Questionnaire. https://sites.google.com/site/theipaq/scoring-protocol. Accessed 6 Mar 2019.
49. Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. Int J Behav Nutr Phys Act. 2011;8:115.
50. Mai KS, Sandbaek A, Borch-Johnsen K, Lauritzen T. Are lifestyle changes achieved after participation in a screening programme for type 2 diabetes? The ADDITION study, Denmark. Diabet Med J Br Diabet Assoc. 2007;24:1121–8.
51. Danquah IH, Petersen CB, Skov SS, Tolstrup JS. Validation of the NPAQ-short – a brief questionnaire to monitor physical activity and compliance with the WHO recommendations. BMC Public Health. 2018;18:601.
52. WHO | Global recommendations on physical activity for health. https://www.who.int/dietphysicalactivity/publications/9789241599979/en/. Accessed 30 Dec 2018.
53. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy Amst Neth. 1990;16:199–208.
54. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20:1727–36.
55. Health-Related Quality of Life Measurement Inpatient Geriatric Rehabilitation: A Comparison of the Feasibility and Suitability Between the SF-36 and EQ-5D-5L Questionnaires| Abstract. https://www.omicsonline.org/peer-reviewed/healthrelated-quality-of-life-measurement-inpatient-geriatric-rehabilitation-a-comparison-of-the-feasibility-and-suitability-betw-97234.html. Accessed 15 Mar 2019.
56. Haywood KL, Garratt AM, Fitzpatrick R. Quality of life in older people: a structured review of generic self-assessed health instruments. Qual Life Res Int J Qual Life Asp Treat Care Rehabil. 2005;14:1651–68.
57. Janssen MF, Pickard AS, Golicki D, Gudex C, Niewada M, Scalone L, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. Qual Life Res. 2013;22:1717–27.
58. EuroQol: Et generisk mål for helbredstilstand - Claire Gudex | bibliotek.dk. https://bibliotek.dk/da/work/870971-tsart:84636797. Accessed 6 Mar 2019.
59. Russell DW. UCLA loneliness scale (version 3): reliability, validity, and factor structure. J Pers Assess. 1996;66:20–40.
60. Lee J, Cagle JG. Validating the 11-item Revised University of California los Angeles scale to assess loneliness among older adults: an evaluation of factor structure and other measurement properties. Am J Geriatr Psychiatry. 2017;25:1173–83.
61. Resnick B, Spellbring AM. Understanding what MOTIVATES older adults to exercise. J Gerontol Nurs. 2000;26:34–42.
62. McAuley E. The role of efficacy cognitions in the prediction of exercise behavior in middle-aged adults. J Behav Med. 1992;15:65–88.
63. Resnick B, Orwig D, Wehren L, Zimmerman S, Simpson M, Magaziner J. The exercise plus program for older women post hip fracture: participant perspectives. The Gerontologist. 2005;45:539–44.
64. Resnick B, Zimmerman S, Orwig D, Furstenberg AL, Magaziner J. Model testing for reliability and validity of the outcome expectations for exercise scale. Nurs Res. 2001;50:293–9.
65. Resnick B, Zimmerman SI, Orwig D, Furstenberg AL, Magaziner J. Outcome expectations for exercise scale: utility and psychometrics. J Gerontol B Psychol Sci Soc Sci. 2000;55:S352–6.
66. Due P, Holstein B, Lund R, Modvig J, Avlund K. Social relations: network, support and relational strain. Soc Sci Med. 1999;48:661–73.
67. Lund R, Mortensen EL, Christensen U, Bruunsgaard H, Holm-Pedersen P, Fiehn N-E, et al. Cohort profile: the Copenhagen aging and midlife biobank (CAMB). Int J Epidemiol. 2016;45:1044–53.
68. Lund R, Nielsen LS, Henriksen PW, Schmidt L, Avlund K, Christensen U. Content validity and reliability of the Copenhagen social relations questionnaire. J Aging Health. 2014;26:128–50.
69. World Health Organization. WHO | What is Moderate-intensity and Vigorous-intensity Physical Activity? WHO; 2017. http://www.who.int/dietphysicalactivity/physical_activity_intensity/en/. Accessed 26 Jun 2017.
70. World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world. 2018.
71. World Health Organization. Global strategy and action plan on ageing and health. 2017. http://www.who.int/ageing/WHO-GSAP-2017.pdf?ua=1.
72. Bandholm T, Christensen R, Thorborg K, Treweek S, Henriksen M. Preparing for what the reporting checklists will not tell you: the PREPARE trial guide for planning clinical research to avoid research waste. Br J Sports Med. 2017;51:1494–501.
73. Nichol AD, Bailey M, Cooper DJ. Challenging issues in randomised controlled trials. Injury. 2010;41:S20–3.
74. Heneghan C, Goldacre B, Mahtani KR. Why clinical trial outcomes fail to translate into benefits for patients. Trials. 2017;18. https://doi.org/10.1186/s13063-017-1870-2.
75. Hansen BH, Dalene KE, Ekelund U, Fagerland MW, Kolle E, Steene-Johannessen J, et al. Step by Step: Association of Device-Measured Daily Steps With All Cause-Mortality – a Prospective Cohort Study. Scand J Med Sci Sports. 2020. https://doi.org/10.1111/sms.13726.
76. Lee I-M, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of Step Volume and Intensity with all-Cause Mortality in older women. JAMA Intern Med. 2019;179:1105–12.
77. Marcus BH, Napolitano MA, King AC, Lewis BA, Whiteley JA, Albrecht A, et al. Telephone versus print delivery of an individualized motivationally tailored physical activity intervention: project STRIDE. Health Psychol. 2007;26:401–9.
78. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c869.

79. Middleton KR, Anton SD, Perri MG. Long-term adherence to health behavior change. Am J Lifestyle Med. 2013;7:395–404.
80. REACH | HORIZON 2020. http://reach2020.eu/?page_id=1485. Accessed 23 Nov 2017.

Publisher's Note

APPENDIX 6: PAPER F (STUDY IV)

Title: The MIPAM trial – Motivational Interviewing and Physical Activity Monitoring to enhance the daily Level of Physical Activity among Older Adults – a Randomized Controlled Trial

Authors:
Rasmus Tolstrup Larsen, PT, MSc [1]
Christoffer Bruun Korfitsen, PT, MSc [1, 2]
Camilla Keller, PT [3]
Jan Christensen, PT, MSc, PhD [4]
Henning Boje Andersen [5], Professor, Technical University of Denmark
Carsten Juhl [6] PT, MPH, PhD
Henning Langberg [7] PT, PhD, DMSc Professor

Affiliations:
1: Section of Social Medicine, Department of Public Health, University of Copenhagen, Denmark
2: Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Denmark
3: Research Governance, Evaluation & Communication, Danish Cancer Society Research Center, Copenhagen, Denmark
4: Department of Occupational- and Physiotherapy, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark
5: Technical University of Denmark, Department of Technology, Management and Economics
6: Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark
Department of Physiotherapy and Occupational Therapy, Copenhagen University Hospital, Herlev and Gentofte, Denmark
7: Section of Health Services Research, Department of Public Health, University of Copenhagen, Denmark

Key words: Older Adults, Elderly, Community-Dwelling, Health technology, Physical activity, Monitoring, Motivational Interviewing

Abstract

Background

One in four older adults in Denmark and almost half of the very old above 75 do not meet the World Health Organization's recommendations for a minimum of physical activity (PA). A cost-efficient and effective way to increase focus on and motivation for daily walking might be to use Physical Activity Monitors (PAMs) in combination with behavioural change intervention. Thus, the objective of this randomized controlled study was to investigate the effect of Motivational Interviewing (MI) as an add-on intervention to a PAM-based intervention measured in community-dwelling older adults.

Methods

This two-arm parallel group randomized controlled trial compared a 12-weeks PAM-based intervention with additional MI (PAM+MI group) with a PAM-based intervention alone (PAM group). The primary outcome, average daily step count, was analysed with a linear regression model, adjusted for sex and baseline daily step count. Following the intention-to-treat principle, multiple imputation based on baseline step count, sex and age was performed.

Results

In total, 38 participants were randomized to the PAM intervention and 32 to the PAM+MI intervention arm. During the intervention period, PAM+MI participants walked on average 909 more steps per day than PAM participants, however insignificant (95%CI: -71; 1889) and reported 2.3 points less on the UCLA Loneliness Scale (95%CI: -4.5; -1.24).

Conclusion

The use of MI, in addition to a PAM-based intervention among older adults in PA promoting interventions hold a potential clinically relevant effect on physical activity and should thus be investigated further with adequately powered RCTs.

Introduction

Background and objectives

More than 50 percent of the European older adults are insufficiently physically active [1]. Higher levels of physical activity (PA) among older adults are associated with positive health-related outcomes, including lower levels of frailty [2] and lower levels of all-cause mortality [3]. Furthermore, inactivity among [1] older adults are associated with higher levels of non-communicable diseases, lower functional health, higher risk of depression and cognitive decline [4–6]. Thus, physical inactivity is one of the leading causes of major non- communicable diseases [7]. Furthermore, strong evidence also exists on the positive effects of PA on several chronic diseases including dementia, type 2 diabetes, coronary heart disease, chronic obstructive pulmonary disease, osteoarthritis and several cancers [8].

The *Global Action Plan on Physical Activity "More active People for a healthier World"* published by the WHO in 2018 states: *"global progress to increase physical activity has been slow, largely due to lack of awareness and investment"* [9]. Especially in older adults, easy access to effective PA programs can benefit societies by allowing older adults to maintain an active life and independent living [9]. As walking has been shown to be the most frequent PA modality among older adults [10] and daily step counts to be highly associated with all-cause mortality and cardiovascular disease-morbidity [11], large scale PA programs should include focus on increasing the level of walking in exercising and the amount of walking in ambulant activities.

A cost-efficient way to increase focus on and motivation for a higher level of daily walking is to use Physical Activity Monitors (PAMs) in PA interventions among older adults. A recent systematic review and meta-analysis concluded that the use of feedback from PAMs among older adults was safe, feasible and moderately effective, equivalent to an additional 1300 daily steps, in increasing the daily level of PA [12,13]. The Internet-Of-Things and wearables in medicine are here to stay [14,15], and future studies should not investigate effectiveness from the PAMs themselves, but use active comparisons to clarify how Behaviour Change Theories (BCTs) can support wearable devices and self-monitoring of behaviour [12,13,16,17].

Self-monitoring, goal setting, action planning, information about behaviour-health links and the consequences of inactivity are important BCTs in PA-interventions [17–21]. Motivational Interviewing

(MI) guides the participants using empathic listening, self-reflection and counselling [22], and aims to facilitate positive behavioural change through increased motivation and increased self-efficacy [23,24]. MI alone has been shown to be short-term effective in increasing PA among older adults with heart failure [25] and hip fracture [26]. Furthermore, older adults have found the combination of MI and PAM-interventions acceptable [27].

While passive comparisons with PAM-based interventions are no longer needed, clarification on the effectiveness of PAM-based interventions in combination with BCT-interventions is needed [12,13]. Thus, the objective of this study was to investigate the short-term effect of MI as an add-on intervention to a PAM-based intervention on average daily step count in community-dwelling older adults.

Methods

Trial design

The MIPAM trial was conducted as a 12-week, investigator-blinded, two-arm parallel-group, superiority randomized controlled trial. This manuscript has been reported according to the CONsolidated Standards of Reporting Trials (CONSORT) 2010 guideline [28]. The allocation ratio between the groups was 1:1 and the only changes to the study protocol [29] was the inability in reaching the desired sample size in the available time period. The methods of this study are described in detail in the study protocol [29].

Ethics

The National Committee on Health Research Ethics informed the authors that the trial, being a non-invasive intervention, is not subject to the Danish laws on research ethics (Journal-nr.:18004960). The plan for managing personal and health information of the trial was approved by The Danish Data Protection Agency (Reference number: 514-0268/18-3000). Prior to agreeing and signing the consent survey, the participants received written information about the study. Informed consent from the participants was collected electronically before filling out the baseline questionnaire.

Participants

Participants were considered eligible for inclusion if they: 1) were retired from the labour market and community-dwelling, 2) were at least 70 years old by the day of enrolling the trial, 3) owned a smartphone or tablet able to install the *Garmin Connect application,* 4) had an active e-mail address 5) were able to fulfil the electronic study survey, and 6) had hearing abilities sufficient to receive oral information about the study and to receive a telephone-based MI intervention. The retirement age in Denmark is currently 65.5 years and is gradually increasing. The age criterion of 70 years was used to avoid including participants between 65 and 70 years who are still fully or partially employed and thus to increase generalizability.

Participants were excluded, if the : 1) had cognitive impairment or mild to severe dementia, 2) were undergoing active chemotherapy or palliative care for cancer, or 3) had a major mobility impairment preventing them from walking.

Interventions

The PAM group received a PAM-based PA promoting intervention and the PAM+MI group received the PAM-based PA promoting intervention and an MI-intervention as an add-on intervention.

Physical activity monitor intervention (PAM)

Participants received a PAM for everyday use in the intervention period and a pamphlet with the national recommendations on PA in aging populations. The specific PAM used in this study is the hip-worn Garmin Vivofit 3 device linked to a pre-specified Garmin Connect account set up with an automatically adjusting daily goal-setting. The participants were asked to wear the PAM for all waking hours, except when bathing, every day for the 12-week intervention period. Participants who experienced installation difficulties received telephone support from the research team not including the blinded primary investigator.

Physical activity monitor intervention plus Motivational interviewing (PAM+MI)

The experimental intervention consisted of the PAM intervention in combination with an MI intervention. During the 12-week intervention period, the participants were scheduled to receive seven telephone calls from trained and certified MI-counsellors in intervention week 1, 2, 3, 5, 7, 9 and 12. The MI-intervention was person-centred and participants were guided with self-reflective counselling and received feedback on their health behaviours in relation to the national recommendations [22]. The Social Cognitive Theory and The Transtheoretical Mode were the theoretical frameworks that guided the intervention content to each individual [30,31]. Self-efficacy and outcome expectations are key constructs and are, among other factors, significant predictors of PA behaviours [31]. Self-efficacy, in this setting for exercise, was operationalized by facilitating confidence when facing barriers to PA, self-monitoring including behavioural goal setting and action planning. Outcome expectancies was operationalized by providing information about behaviour-health link, providing information about consequence and discussion of benefits of and barriers to health behavioural change, which should lead to increased perception of benefits and decreased perception of barriers. Social support was operationalized by identification of supports for maintenance of health behavioural change, and specific goal setting for using supports, which should lead to increase level of support for the participant's health behavioural change.

In this study, participants in the PAM+MI group were encouraged to use a variety of significant supports including family and friends, as well as neighbourhood and specific community resources (e.g., walking groups proposed by the MI-counsellor).

Fidelity

The project MI counsellors were physiotherapists with additional training and education in the MI approach to telephone-based health behaviour counselling. During the study, with participants' verbal consent, telephone MI sessions were audiotaped on a regular basis to ensure fidelity of intervention delivery and to provide counsellor feedback. Based on a review of these recordings a random segment of 20 minutes was selected for rating with the Motivational Interviewing Treatment Integrity Scale version 4 (MITI 4) [32], by two independent coders. The MITI 4 is a reliable measure of proficiency in MI practice as defined by Moyers et al [32]. The MITI 4 consists of four global ratings (Cultivating Change talk, Softening Sustain Talk, Partnership, and Empathy), which are scored on a Likert-type scale from 1 (low) to 5 (high), and 10 individual behaviour counts (Questions, Simple Reflections, Complex Reflections, Persuade with Permission, Giving Information, Affirmations, Emphasize Autonomy, Seeking Collaboration, Persuade and Confront), which are counted within the time frame of the interview [32]. The MI-coders individually coded and reached consensus on MI-behaviour. A median global score in each domain of 4 and a Reflection to Question ratio of >1 were considered adequate MI proficiency.

Outcomes

Primary Outcome Measure:

The average number of steps per day throughout the 12-week intervention period, measured daily and objectively by the hip-worn Garmin Vivofit 3 tri-axial accelerometer, was the primary study outcome. The Garmin Vivofit 3 has been validated along with three other monitors and the hip-worn PAMs were found to be superior to wrist-worn PAMs in terms of measurement properties among older adults with and without rollators [33].

Secondary Outcome Measures:

Secondary outcome measures included self-reported information from the participants on PA, health-related quality of life, loneliness, self-efficacy for exercise, outcome expectancy for exercise, and social relations. According to the protocol the categories MVPA, walking time and sedentary time were estimated with *The International Physical Activity Questionnaire-Short Form (*IPAQ-SF*)* [34–37]*,* MVPA was estimated with *The Nordic Physical Activity Questionnaire short (*NPAQ-short*)* [38,39]*,* the HRQoL score (EQVAS) was estimated with *The EuroQol-5 Domain (*EQ-5D-5L*) Quality of life questionnaire* [40–43]*,* the total score was estimated with *The UCLA Loneliness Scale* [44,45] to measure loneliness*,* the sum score was from *the Self-Efficacy for Exercise (SEE-DK)* [46] to used to measure self-efficacy*,* and the sum score was from *the Outcome Expectancy for Exercise-2 (OEE2-DK)* [47] to used to measure outcome-expectancy . Secondary outcome measures are described in greater detail in the study protocol [29]. *The Copenhagen Social Relations Questionnaire* (CRSQ) [48,49] was used only to inform the MI-counsellors and to determine whether the participants lived alone.

All secondary outcomes were collected at baseline and at post-intervention. The baseline measurement took place before randomization and thus before the PAM+MI group received their first motivational interview; the post-intervention questionnaire was distributed immediately after the 12-weeks of intervention.

Sample size

To show a moderate effect difference (0.5*standard deviation between group difference) with 80% power and a 0.05 significance level, 128 completed participants were needed. To account for attrition, a 20 percent dropout was expected and thus, 154 participants (77 in each group) were needed to be allocated to each of the two groups.

Randomization

Participants were randomly assigned to either the intervention or the PAM group, with a 1:1 allocation ratio. Eligible participants who completed the baseline period of one week were randomized in blocks of minimum four participants, stratified on sex and average daily baseline step count for the baseline period. STATA statistical software was used to conduct the stratified randomization. Allocation was concealed for the primary investigator. One investigator (JC) was responsible for the randomization process and had no role in the recruitment of participants nor in the statistical analyses.

Blinding

The primary investigator (RTL), who was responsible for analyses and data-management, was blinded for participant allocation until the last participant completed the post-intervention questionnaire. As the secondary outcome measures are self-reported, outcome assessor cannot be considered blinded. Due to the nature of the intervention neither participants nor physiotherapists conducting the motivational interviews could be blinded to allocation.

Data collection and management

Information about data collection management can be found in the study protocol. No deviations from the protocol occurred on this matter [29].

Statistical methods

Distributions of continuous data was evaluated by inspecting Quantile-Quantile plots of the standardized residuals and histograms with normal distribution curves. Continuous data with normal distributions was analysed with parametrical statistics and summarized with means and standard deviations. Continuous data without normal distribution was analysed as ordinal data with

non-parametrical statistics and summarized with medians and interquartile ranges. Categorical or binary data were summarized with frequencies and percentage of total.
The primary outcome, average daily step count, was analysed with a linear regression model investigating the between-group differences, adjusted for sex and baseline daily step count. Following the intention-to-treat principle, Gaussian normal regression method with predictive mean matching was used to impute missing values (multiple imputation based on baseline step count, sex and age) where less than seven days of step counts were available for the intervention period. The same procedure was used to analyse between group differences on secondary outcomes. All secondary outcomes were adjusted for baseline score of the specific outcome, baseline daily step count and sex. Harms, as defined in the study protocol [29], were evaluated by calculating the relative risk (RR), separately for serious and non-serious adverse event between the intervention and PAM group [50]. In calculating the average daily step count, days with less than 100 steps were handled as "days of non-wear" and excluded. A post-hoc power calculation was performed with number of participants, effect size of the between group difference from the primary analysis on daily steps and the baseline overall standard deviation on daily steps.

RStudio version 1.1.463 for Mac OS X was used for all statistical analyses and illustrations [51].
The CRAN 'mice' package was used to perform the predictive mean matching multiple imputations and the 'ggplot2' package was used to generate a scatterplot with means and error bars for daily steps throughout the intervention and box plots of secondary outcomes. An alpha level on 0.05 was considered the threshold for statistical significance.

Results

Participant flow and information on discontinued participants

Between May 1, 2019 and January 4, 2020, 79 participants were considered eligible for inclusion and received the trial content. After nine eligible participants refused to participate, 70 participants were included and randomized to one of the two intervention. Of these, 38 were allocated to the PAM intervention arm and 32 to the MIPAM intervention arm. In the PAM intervention arm, 34 participants completed the 12 weeks and four participants discontinued (figure 1). In the MIPAM intervention arm, 28 participants completed the 12 weeks and four persons discontinued their participation (figure 1). Due to low inclusion rate and insufficient funding to extend the inclusion period, it was decided, to stop inclusion of participants to the trial in January 2020. This resulted in an underpowered trial that did not reach the desired sample size of 128 participants excluding dropouts.

CONSORT 2010 Flow Diagram for the MIPAM trial

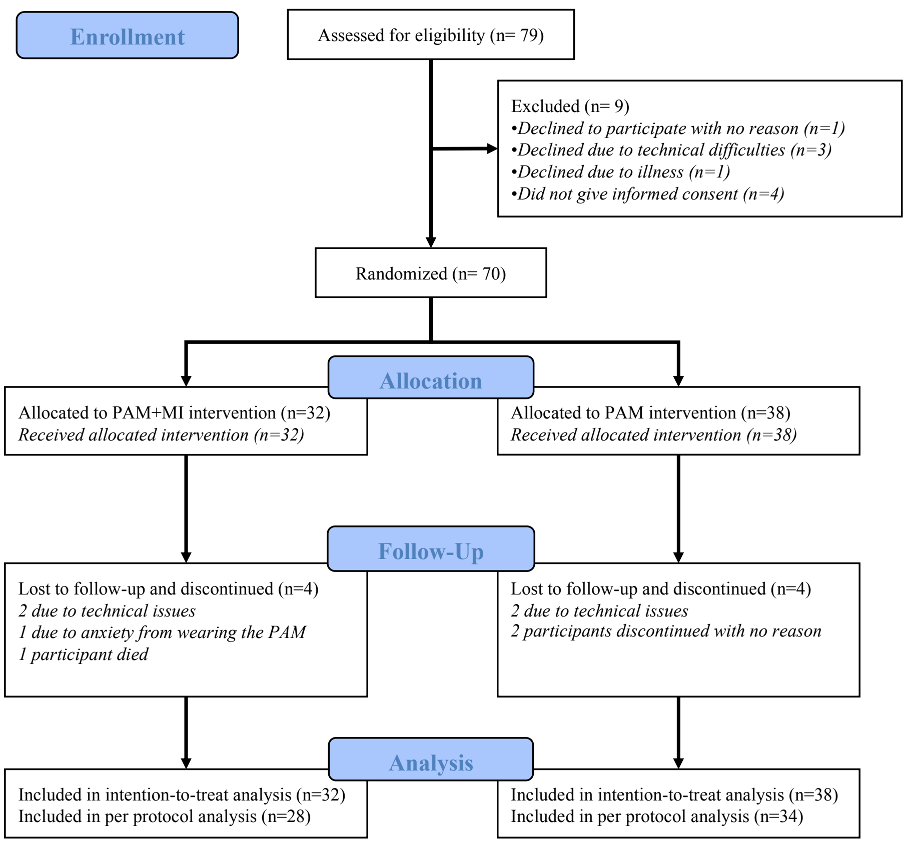


Figure 1. Consort flow diagram.

Baseline data

Socio-demographics and PA characteristics of included participants are reported in table 1. There were no between-group differences on any variables, except for a higher rate of participants in the PAM+MI group reporting pain (51.6% vs 25.0%). The median age of the participants was 72 years, 28 of 70 participants were male (40.0%), 22 reported to have used a PAM before (32.8%) and the mean baseline daily step count was 5,881.

Table 1. Socio-demographics and physical activity characteristics of included participants

	Overall (n=70)	PAM group (n=38)	PAM+MI group (n=32)	*p*
Age, median [IQR]	72.0 [70.0, 74.0]	71.0 [70.0, 74.3]	73.0 [71.0, 74.0]	0.134
Sex, n male (%)	28 (40.0)	16 (42.1)	12 (37.5)	0.613
BMI, mean (SD)	27.2 (4.4)	27.3 (4.9)	27.1 (3.9)	0.581
Education, n (%)				0.522
No education	1 (1.5)	1 (2.8)	0 (0.0)	
Upper secondary education	11 (16.4)	5 (13.9)	6 (19.4)	
Bachelor's degree or equivalent tertiary education level	39 (58.2)	21 (58.3)	18 (58.1)	
Master's degree, equivalent tertiary education level, or above	16 (23.9)	9 (25.0)	7 (22.6)	
Lives alone, n (%)	26 (39.4)	13 (36.1)	13 (43.3)	0.507
Smoking, n (%)				0.509
Never smoked	33 (49.3)	19 (52.8)	14 (45.2)	
Stopped smoking	29 (43.3)	14 (38.9)	15 (48.4)	
Smokes	5 (7.5)	3 (8.3)	2 (6.5)	
In pain, n (%)	25 (37.3)	9 (25.0)	16 (51.6)	0.046
Long-term chronic disease or disability, n (%)	33 (49.3)	16 (44.4)	17 (54.8)	0.379
Limited in usual activities due to disability, health or pain (%)				0.388
Not limited	35 (52.2)	20 (55.6)	15 (48.4)	
Limited to some extend	26 (38.8)	14 (38.9)	12 (38.7)	
Seriously limited	6 (9.0)	2 (5.6)	4 (12.9)	
Walking aids (%)				0.253
None	65 (97.0)	35 (97.2)	30 (96.8)	
Cane	1 (1.5)	0 (0.0)	1 (3.2)	
Rollator	1 (1.5)	1 (2.8)	0 (0.0)	
% of total activity from walking, median [IQR]	69.5 [30.8, 80.0]	64.0 [20.0, 80.0]	69.5 [40.0, 79.5]	0.363
Wants to be more physically active, n (%)				0.259
Yes	56 (83.6)	28 (77.8)	28 (90.3)	
No	3 (4.5)	2 (5.6)	1 (3.2)	
Do not know	8 (11.9)	6 (16.7)	2 (6.5)	
Have used or uses physical activity monitor, n (%)	22 (32.8)	12 (33.3)	10 (32.3)	0.997
UCLA Loneliness Scale Sum, mean (SD)	32.9 (8.6)	33.5 (9.5)	32.3 (7.5)	0.399
EQ-5D-5L				
Problems with mobility, n (%)	27 (40.1)	13 (36.1)	14 (46.7)	0.373
Problems with self-care, n (%)	4 (6.1)	2 (5.6)	2 (6.7)	0.995
Problems with usual activities, n (%)	19 (28.8)	9 (25.0)	10 (33.3)	0.442
Problems with pain and discomfort, n (%)	43 (65.2)	20 (55.6)	23 (76.7)	0.087
Problems with anxiety and depression, n (%)	13 (19.7)	7 (19.4)	6 (20.0)	0.995
EQ Visual Analogue Scale, median [IQR]	80.0 [70.0, 90.0]	85.0 [70.0, 90.0]	80.0 [70.0, 90.0]	0.438
Outcome Expectancy for Exercise-2 Scale Sum, mean (SD)	51.6 (6.9)	50.3 (7.27)	53.1 (6.1)	0.074
Self-Efficacy for Exercise Scale Sum, mean (SD)	60.5 (19.8)	59.4 (20.15)	61.8 (20.0)	0.442
Baseline steps per day, mean (SD)	5881 (2948)	6029 (3009)	5705 (2913)	0.649
International Physical Activity Questionnaire Short Form				
Minutes of vigorous activity per day, median [IQR]	0.0 [0.0, 24.1]	0.0 [0.0, 19.3]	0.0 [0.0, 24.1]	0.581
Minutes of moderate activity per day, median [IQR]	0.0 [0.0, 24.1]	0.0 [0.0, 19.3]	0.0 [0.0, 24.1]	0.581
Minutes of MVPA per day, median [IQR]	0.0 [0.0, 48.2]	0.0 [0.0, 38.6]	0.0 [0.0, 48.2]	0.581
Minutes of walking activity per day, median [IQR]	27.9 [0.0, 327.9]	23.6 [0.0, 422.1]	31.1 [0.0, 214.3]	0.643
MET-minutes per day, median [IQR]	316.9 [63.3, 1386.0]	394.1 [40.9, 1432.0]	254.6 [102.5, 850.5]	0.570
Minutes of sedentary time per day, median [IQR]	303.0 [210.0, 480.0]	303.0 [240.0, 435.0]	316.0 [202.5, 480.8]	0.676
IPAQ physical activity categories, n (%)				0.690
High	18 (27.3)	10 (27.8)	8 (26.7)	
Low	26 (39.4)	14 (38.9)	12 (40.0)	
Moderate	22 (33.3)	12 (33.3)	10 (33.3)	

Nordic Physical Activity Questionnaire short				
Minutes of MVPA per day, median [IQR]	30.0 [12.9, 68.6]	38.6 [8.2, 82.5]	25.7 [17.5, 46.1]	0.227
Minutes of moderate activity per day, median [IQR]	11.4 [0.0, 24.6]	7.9 [0.0, 34.3]	12.1 [0.0, 19.8]	0.651
Minutes of vigorous activity per day, median [IQR]	17.1 [0.0, 34.3]	19.3 [0.0, 58.4]	13.6 [0.0, 21.4]	0.225
NPAQ physical activity categories, n (%)				0.231
Inactive	8 (12.1)	4 (11.1)	4 (13.3)	
Insufficiently physically active	10 (15.2)	7 (19.4)	3 (10.0)	
Sufficiently physically active	12 (18.2)	4 (11.1)	8 (26.7)	
Optimally physically active	36 (54.5)	21 (58.3)	15 (50.0)	
NPAQ compliance with WHO recommendations, n (%)	48 (72.7)	25 (69.4)	23 (76.7)	0.490

BMI: Body Mass Index, PAM: Physical Activity Monitor, EQ-5D: EuroQol Research Foundation Five Domains, UCLA: University of California Los Angeles, OEE: Outcome Expectancy for Exercise, SEE: Self Efficacy for Exercise. IPAQ-SF: International Physical Activity Questionnaire-Short Form, NPAQ: Nordic Physical Activity Questionnaire-Short Form, MVPA: Moderate to Vigorous Physical Activity, SD: standard deviation, IQR: interquartile range, IPAQ-SF: International Physical Activity Questionnaire-Short Form, NPAQ: Nordic Physical Activity Questionnaire-Short Form, MVPA: Moderate to Vigorous Physical Activity. Test for between-group difference in normal distributed continuous variables (BMI, UCLA Loneliness Scale Sum Score, Outcome Expectancy for Exercise-2 Scale Sum Score, Self-Efficacy for Exercise Sum Score and Baseline Daily Steps) were performed with unpaired t-test, test for between group difference in non-normal distributed continuous (age, % of total activity from walking, EQ Visual Analogue Scale, all IPAQ and NPAQ scores) variables were performed with Mann-Whitney U test, test for between group difference in categorical or binary variables with Chi2 test, p-values ≤ 0.05 are considered significant.

Numbers analysed

The median days of missing PA data during the 12 weeks of intervention was 6 [IQR: 1, 32] days in the PAM group and 4.5 [IQR: 0.75, 26] in the PAM+MI group. Data for four participants were imputed for average daily steps. Data for six participants were imputed for IPAQ-SF MVPA and minutes of sedentary time per day, NPAQ-Short MVPA minutes per day, EQ-VAS, UCLA Loneliness Scale Sum Score, and SEE-DK Sum Score. Data for seven participants were imputed for IPAQ-Short minutes of walking per day and OEE2-DK Sum.

Outcomes and estimation

For the primary outcome, the PAM+MI group increased by 909 steps daily throughout the intervention period compared to the PAM group, but insignificantly (95%CI: -71; 1889). For the secondary outcomes, the participants in the PAM+MI group reported 2.3 UCLA Loneliness Scale Sum Score points less compared to the PAM group (95%CI: -4.5; -1.2). No relevant or significant differences were found in the other secondary outcomes.

Table 2. Results from multiple regression models on outcomes

Outcome	**Post-intervention scores**		Adjusted between group difference from multiple regression model		
	PAM group (n=38) mean (95%CI)	**PAM+MI group (n=32)** mean (95%CI)	**Between group difference**	**95%CI**	***p***
Average daily steps	5,837 (4,932; 6,742)	6,492 (5,472; 7,513)	909	(-71; 1889)	0.07
IPAQ-SF					
MVPA minutes per day	53.9 (15.3; 92.5)	34.4 (5.2; 63.6)	-0.2	(-46.3; 45.8)	0.992
Minutes of walking per day	149.2 (59.1; 239.3)	218.5 (111.5; 325.5)	78.1	(-6.1; 217.3)	0.266
Minutes of sedentary time per day	358.5 (303.6; 413.4)	335.0 (273.0; 397.0)	-40.3	(-102.8; 22.1)	0.201
NPAQ-Short					
MVPA minutes per day	72.5 (41.0; 104.0)	66.6 (40.1; 93.1)	-3.8	(-45.3; 37.7)	0.856
EQ-VAS	80.6 (76.0; 85.1)	81.6 (78.2; 85.1)	2.9	(-1.9; 7.7)	0.227
UCLA Loneliness Scale Sum Score	32.8 (29.6; 36.0)	30.2 (27.4; 33.0)	-2.3	(-4.5; -1.2)	0.04
Self-Efficacy for Exercise Sum Score	52.5 (45.9; 59.1)	55.3 (45.9; 60.4)	3.5	(-4.3; 11.2)	0.375
Outcome Expectancy for Exercise-2 Sum Score	51.3 (48.5; 54.2)	53.2 (50.5; 56.0)	2.0	(-2.0; 6.0)	0.320

Abbreviations: IPAQ-SF: International Physical Activity Questionnaire Short Form, NPAQ-Short: Nordic Physical Activity Questionnaire Short, EQ-VAS: EuroQol Visual Analogue Scale, UCLA: University of California, Los Angeles, Data for four participants were imputed for average daily steps. Data for six participants was imputed for IPAQ-SF MVPA and minutes of sedentary time per day, NPAQ-Short MVPA minutes per day, EQ-VAS, UCLA Loneliness Scale Sum Score, and SEE Sum Score. Data for seven participants was imputed for IPAQ-Short minutes of walking per day and OEE-2 Sum. End point scores are unadjusted. Primary analysis is the multiple linear regression model adjusted for baseline score, baseline steps, age and sex. Coefficients > 0 means higher value in the PAM+MI group. Negative coefficients for IPAQ-SF Sedentary Time and UCLA Loneliness Scale Sum Score means less sedentary time and loneliness in the PAM+MI group. P-values < 0.05 is considered significant.

Figure 2 illustrates unadjusted steps per day for the two study arms through the study period. In the appendix, figure 3 and figure 4 illustrates box plots of other secondary outcomes at baseline and end point for both treatment arms.

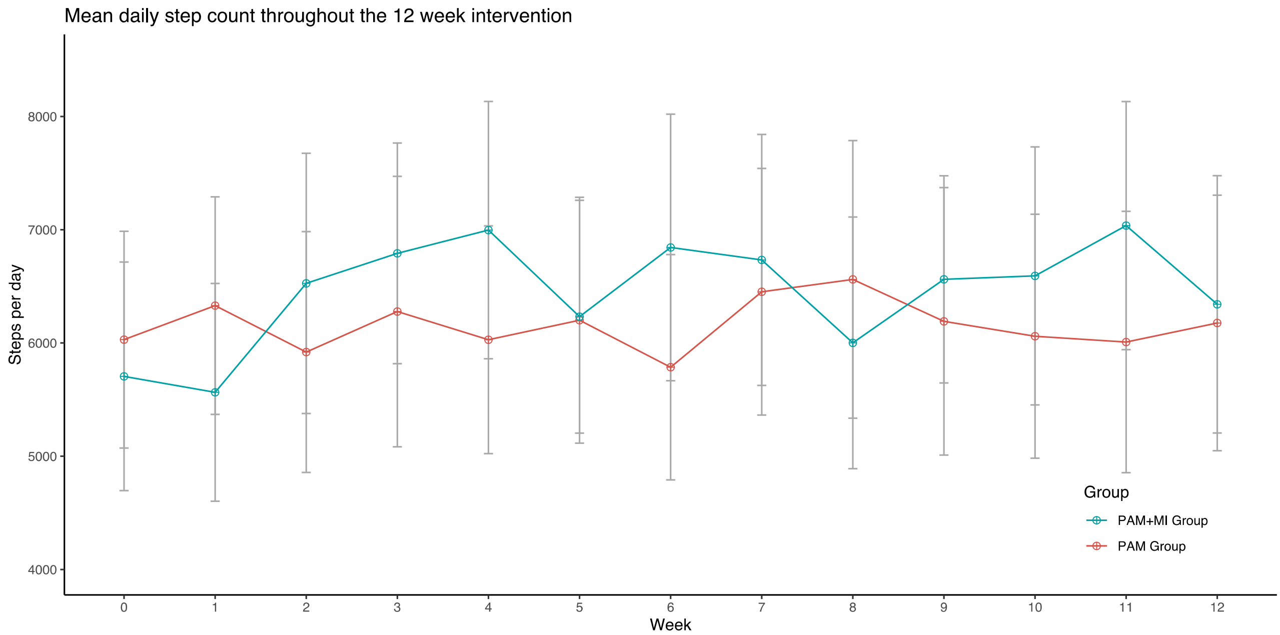


Figure 2. Unadjusted mean daily step counts throughout the 12-week intervention. W0: baseline week. Intervention period: w1 to w12. Circles represent mean values and error bars represent 95% confidence intervals.

Fidelity

Each participant in the PAM+MI group was scheduled to receive seven MI calls. Among the 28 complete case PAM+MI group participants, 23 (82.1%) received all seven calls, four (14.3%) received six calls and one participant (3.6%) received four calls. In total, 170 calls with an average length of 18.4 minutes were delivered to the PAM+MI group. Six MI calls were audiotaped and coded by the two coders using the Motivational Interviewing Treatment Integrity Scale. The median cultivating-change talk global score was 3.5, the median softening sustain talk was 4, the median partnership score was 4, and the median empathy score was 4. The median number of Giving Information was 3.5, the median of Simple Reflections was 3, the median of Questions and Complex Reflections was 7, the median of Affirm and Seek was 1 and 1.5 respectively and the median number of Persuade, Persuade with Permission, Emphasize Autonomy and Confront was 0. The ratio of Reflections to Questions was 1.3.

Ancillary analyses

The eight participants who discontinued the intervention differed significantly from the complete cases as they were older 78.5 years [IQR: 74.0, 81.5] compared with 72.0 years [IQR: 70.0, 74.0], p=0.035, only female (54.8% female in complete case versus 100% female in discontinued participants, p=0.038), and had a different use of walking aids (one rollator user and no cane users in the complete case versus with a cane user and no rollator users in discontinued participants, p=0.006). No other significant or clinically relevant differences were found on other baseline variables.

A post-hoc power calculation of the primary analysis, showed a level of power on 24.6%. This analysis included 70 participants, an effect size on 909 steps, the standard deviation of the daily step count of 2948 and an alpha level of 0.05.

Harms

The frequencies of dropouts from the two groups were similar with four dropouts in the PAM+MI group (12.5%) and four dropouts in the PAM group (10.5%). Two participants, both allocated to the PAM+MI group (6.3%), discontinued due to adverse events, as judged by the investigators. One participant died and one participant had increasing anxiety of wearing the PAM triggering existing mental illness. There was no significant between group difference between the proportions of adverse events in the groups (0% in the PAM group versus 6.3% in the PAM+MI group, p=0.400).

Discussion

To our knowledge, the MIPAM trial is the first study to investigate the effect of adding MI to a PAM-based PA intervention among community-dwelling older adults aged 70 or above. As this study had insufficient power no final conclusions can be drawn about the true effect of the intervention. However, the PAM+MI group walked on average 909 (95%CI: -71; 1889) steps more per day compared to the PAM group. Even though this finding is non-significant, the confidence interval suggest that MI might possible increase the average level of PA when adding it to a PAM-based PA intervention.

The research team chose objectively measured PA as the primary outcome of interest as the aim of trial was to investigate behaviour change related to PA. However, the real-world interest of clinicians and healthcare workers might not be focused on the PA levels among older adults but on hard outcomes such as disease prevalence and mortality. Thus, daily PA might not be categorized as critical for decision making [52] and the results of this trial cannot be extrapolated to conclude upon the associations between the measured behavioural change and critical outcomes. However, PA levels among older adults are associated with levels of non-communicable diseases, functional health, risk of depression and cognitive function [4–6]. As physical inactivity remains one of the leading causes of major non-communicable diseases [7], daily PA levels serve as a highly relevant construct to measure and as one of the most important surrogate outcomes for critical outcomes among older adults [53,54]. Evidence suggests that a PA level of 7,100 steps per day (if averaged over a week) is enough for older adults to meet WHO recommendations for PA [55]. Additionally, for each increment of 1,000 steps per day, the risk of all-cause mortality decreases with 11% even after being adjusted for several confounding factors [56]. In summary, MI might hold the potential of keeping older adults more physically active over a 12-week intervention study, and the difference between the MI plus PAM and the PAM alone group is clinically relevant for older adults.
This trial failed to reach a sample size of 128 participants and consequently should be categorized as underpowered. The post-hoc power calculation revealed a 24.6% power in this specific study, for being able to reject the null hypothesis, when it should be rejected. Thus, the between group difference on 909 (95%CI: -71.; 1889) steps per day may be an overestimation. However, when inspecting the confidence interval for the primary analysis, the between-group difference lies between 71 steps in favour of the PAM group and 1,889 steps in favour of the PAM+MI group.

Hence, it seems very plausible that the PAM+MI group had a higher daily step count in the intervention period.

Physical activity is a difficult construct to measure with many considerations about practicality, feasibility and validity [57]. The existing literature on randomized controlled MI-based studies investigating physical activity in older adults uses different measures of physical activity including objectively measurement of subgroups of the study, comparing accelerometer measured baseline weeks with end-point weeks, recall questionnaires and physical activity diaries [25–27,58,59]. Physical activity is a behaviour that should be measured consequently over the period of interest, especially in intervention research as the changes and between group differences might occur during the trial, and not before and after the trial. This problem also exists in the observational literature linking disease, morbidity and mortality with physical activity in older adults or in general [3,56,60,61], and even though these observational studies includes large samples which leads to precise estimates, the Hawthorne effect, defined as immediate behavioural change expected from research participation, cannot be ruled out and might impose different types of bias [62]. Especially in moderately sized experimental behavioural change studies, physical activity should be measured consecutively and conclusions should be drawn on accumulated or average physical activity and not on point estimates. When inspecting the means of figure 2, it comes clear that a high degree of variability exists from week to week and if one of these weeks were used as end-point outcome alone, different conclusions could be drawn. However, the variability could also be explained by the large variation in the data and the relatively few samples, but as a methodological consideration it comes clear that behavioural change studies should include consecutively measured constructs, which is a strength of this study.

The secondary outcomes assessed in this study include self-reported PA, health related quality of life, loneliness, self-efficacy for exercise and outcome-expectancy for exercise. Besides UCLA Loneliness Scale, none of the secondary outcomes were significantly different between the groups at endpoint. The UCLA Loneliness Scale Sum Score was 2.3 points (95%CI -4.5; -1.24) lower in the PAM+MI group. Because the literature lacks a minimal clinically important differences on the UCLA Loneliness scale among older adults, this can also be interpreted as a small to moderate effect size (Cohens d: 0.38) [63]. However, this can easily be explained by the nature of the intervention, as MI uses active empathic listening, self-reflection and counselling [22], which naturally affects some of the items used in the UCLA Loneliness scale. Furthermore, some of the difference can also be explained by a small insignificant difference between the groups at baseline and extrapolation of these results should not be done on this secondary outcome, but on future well-powered studies using Loneliness as the primary outcome. In summary, with all the limitations to the finding on loneliness, it is still a relevant difference, and as loneliness has been reported to affect self-reported health and PA negatively this finding might be associated with a higher activity level in the PAM+MI group [64].

To our knowledge, most studies published on MI interventions among older adults targeting PA behaviour directly include intervention lengths from 8 weeks to 6 months [25–27,59]. Furthermore, a systematic review and meta-analysis reported the median length of PAM-based interventions among older adults to be 12 weeks, ranging from 4 to 52 weeks [12]. Hence, the intervention length on 12 weeks of this study is in line with former studies. It is possible that the exposure to MI was too short to demonstrate an actual effect within the 12 weeks and that the results of this trial only reflects the initial and short-term behavioural changes and thus not the long-term effects. To investigate the long-term effects of this 12-week intervention, 6 and 12-month follow-ups will be conducted as it is hypothesized that the MI intervention will help the participants develop more effective strategies to ensure long-term adherence to healthy PA behaviour. Even so, the quality of the MI calls was considered adequate as our results on global ratings of the content ranged from 3.5 to 4 out of 5 (where higher scores indicate higher integrity of the content) and the reflection to question ratio was 1.3. As previously described, MI has been reported to be effective on short-term outcomes, but a recently published study with 1,742 participants did not report any effect of either group-based or individual-based MI [65]. Thus, MI might be effective in some populations, and not in others.

Limitations

This study and the results come with several limitations. Firstly, as previously discussed, the sample size was not large enough to ensure adequate power for this study. Secondly, the inclusion and exclusion criteria for this trial were only ensured by telephone or e-mail by the primary investigator. Thus, it is possible for participants to fall under one or more of the exclusion criteria and still participate in the study, if they (willingly or unwillingly) withheld pertinent information from the primary investigator. Nevertheless, even though this is possible, this potential problem should be balanced between groups, as the randomization occurred after the baseline period. Thirdly, the study participants were not blinded for group allocation and consequently performance bias could have been introduced. This type of problem is common in PA intervention studies, and might cause for an exaggeration of study effects [66]. However, a more recent published meta-analysis of more than 1,100 trials reports found no evidence for an average difference in effect sizes between adequately blinded studies and studies that lack blinding of either participants, healthcare providers or outcome assessors [67]. This trial tried to control for this by using an objectively measured primary outcome that neither participants nor healthcare providers could affect.

Interpretation and reflections from the motivational sessions

Limitations related to the MI-sessions also exists. Reflections from these MI-sessions were not systematically collected and thus should only be used for researchers and health care workers planning to conduct MI among older adults.

Firstly, the first telephone calls were mainly used to form the relationship between the counsellor and the participant and rarely for actual MI-content. Secondly, the participants included in this trial were mainly well-educated, active and resourceful older adults with high levels of health literacy, which might affect the generalizability to the background population of older adults, as previous research has shown that exercise and physical activity adherence are associated with resources such as social support and the ability to understand the benefits of physical activity [68–70].

Lastly, using the Garmin application or navigating the smartphone in general, were frustrating to many participants, however, feedback related to the automatic goal-setting was useful for many. In general, participants were motivated to be committed and pushed to plan more challenging goals, thus these sessions were primarily coaching rather than MI that are normally used among less motivated individuals.

Conclusion

This RCT found a clinically relevant but insignificant difference of 909 (95%CI: -71.; 1889) daily steps in favour of the PAM+MI group. The use of MI, in addition to a PAM intervention, among older adults in PA promoting interventions should be investigated further in sufficiently powered RCTs.

References

1. Lübs L, Peplies J, Drell C, Bammann K. Cross-sectional and longitudinal factors influencing physical activity of 65 to 75-year-olds: a pan European cohort study based on the survey of health, ageing and retirement in Europe (SHARE). *BMC Geriatr*. 2018;18(1):94. doi:10.1186/s12877-018-0781-8

2. Song J, Lindquist LA, Chang RW, et al. Sedentary Behavior as a Risk Factor for Physical Frailty Independent of Moderate Activity: Results From the Osteoarthritis Initiative. *Am J Public Health*. 2015;105(7):1439-1445. doi:10.2105/AJPH.2014.302540

3. Hansen BH, Dalene KE, Ekelund U, et al. Step by Step: Association of Device-Measured Daily Steps With All Cause-Mortality – a Prospective Cohort Study. *Scand J Med Sci Sports*. 2020;n/a(n/a). doi:10.1111/sms.13726

4. WHO | The Global Strategy on Diet, Physical Activity and Health (DPAS). WHO. Accessed January 17, 2019. http://www.who.int/nmh/wha/59/dpas/en/

5. Paterson DH, Warburton DE. Physical activity and functional limitations in older adults: a systematic review related to Canada's Physical Activity Guidelines. *Int J Behav Nutr Phys Act*. 2010;7:38. doi:10.1186/1479-5868-7-38

6. Paterson DH, Jones GR, Rice CL. Ageing and physical activity: evidence to develop exercise recommendations for older adults. *Can J Public Health Rev Can Sante Publique*. 2007;98 Suppl 2:S69-108.

7. Lee I-M, Shiroma EJ, Lobelo F, et al. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet Lond Engl*. 2012;380(9838):219-229. doi:10.1016/S0140-6736(12)61031-9

8. Pedersen BK, Saltin B. Exercise as medicine – evidence for prescribing exercise as therapy in 26 different chronic diseases. *Scand J Med Sci Sports*. 2015;25:1-72. doi:10.1111/sms.12581

9. World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world. Published online 2018.

10. Szanton SL, Walker RK, Roberts L, et al. Older adults' favorite activities are resoundingly active: Findings from the NHATS study. *Geriatr Nur (Lond)*. 2015;36(2):131-135. doi:10.1016/j.gerinurse.2014.12.008

11. Hall KS, Hyde ET, Bassett DR, et al. Systematic review of the prospective association of daily step counts with risk of mortality, cardiovascular disease, and dysglycemia. *Int J Behav Nutr Phys Act*. 2020;17(1):78. doi:10.1186/s12966-020-00978-9

12. Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance amount of physical activity in older adults – a systematic review and meta-analysis. *Eur Rev Aging Phys Act*. 2019;16(1):7. doi:10.1186/s11556-019-0213-6

13. Larsen RT, Christensen J, Juhl CB, Andersen HB, Langberg H. Physical activity monitors to enhance the daily amount of physical activity in elderly—a protocol for a systematic review and meta-analysis. *Syst Rev*. 2018;7(1):69. doi:10.1186/s13643-018-0733-6

14. Dimitrov DV. Medical Internet of Things and Big Data in Healthcare. *Healthc Inform Res*. 2016;22(3):156-163. doi:10.4258/hir.2016.22.3.156

15. Wright SP, Hall Brown TS, Collier SR, Sandberg K. How consumer physical activity monitors could transform human physiology research. *Am J Physiol - Regul Integr Comp Physiol*. 2017;312(3):R358-R367. doi:10.1152/ajpregu.00349.2016

16. Michie S, Richardson M, Johnston M, et al. The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques: Building an International Consensus for the Reporting of Behavior Change Interventions. *Ann Behav Med*. 2013;46(1):81-95. doi:10.1007/s12160-013-9486-6

17. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: A meta-regression. *Health Psychol*. 2009;28(6):690-701. doi:10.1037/a0016136

18. Victor CR, Rogers A, Woodcock A, et al. What factors support older people to increase their physical activity levels? An exploratory analysis of the experiences of PACE-Lift trial participants. *Arch Gerontol Geriatr*. 2016;67:1-6. doi:10.1016/j.archger.2016.06.006

19. Bull ER, McCleary N, Li X, Dombrowski SU, Dusseldorp E, Johnston M. Interventions to Promote Healthy Eating, Physical Activity and Smoking in Low-Income Groups: a Systematic Review with Meta-Analysis of Behavior Change Techniques and Delivery/Context. *Int J Behav Med*. 2018;25(6):605-616. doi:10.1007/s12529-018-9734-z

20. Dusseldorp E, van Genugten L, van Buuren S, Verheijden MW, van Empelen P. Combinations of techniques that effectively change health behavior: Evidence from Meta-CART analysis. *Health Psychol*. 2014;33(12):1530-1540. doi:10.1037/hea0000018

21. Knittle K, Nurmi J, Crutzen R, Hankonen N, Beattie M, Dombrowski SU. How can interventions increase motivation for physical activity? A systematic review and meta-analysis. *Health Psychol Rev*. 2018;12(3):211-230. doi:10.1080/17437199.2018.1435299

22. Miller WR, Rollnick S. *Motivational Interviewing: Helping People Change, 3rd Edition*. Guilford Press; 2013.

23. Cummings SM, Cooper RL, Cassie KM. Motivational Interviewing to Affect Behavioral Change in Older Adults. *Res Soc Work Pract*. 2009;19(2):195-204. doi:10.1177/1049731508320216

24. Hardcastle SJ, Hancox J, Hattar A, Maxwell-Smith C, Thøgersen-Ntoumani C, Hagger MS. Motivating the unmotivated: how can health behavior be changed in those unwilling to change? *Front Psychol*. 2015;6. doi:10.3389/fpsyg.2015.00835

25. Brodie DA, Inoue A. Motivational interviewing to promote physical activity for people with chronic heart failure. *J Adv Nurs*. 2005;50(5):518-527. doi:10.1111/j.1365-2648.2005.03422.x

26. O'Halloran PD, Shields N, Blackstock F, Wintle E, Taylor NF. Motivational interviewing increases physical activity and self-efficacy in people living in the community after hip fracture: a randomized controlled trial. *Clin Rehabil*. 2016;30(11):1108-1119. doi:10.1177/0269215515617814

27. Audsley S, Kendrick D, Logan P, Jones M, Orton E. A randomised feasibility study assessing an intervention to keep adults physically active after falls management exercise programmes end. *Pilot Feasibility Stud*. 2020;6(1):37. doi:10.1186/s40814-020-00570-9

28. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c869. doi:10.1136/bmj.c869

29. Larsen RT, Korfitsen CB, Juhl CB, Andersen HB, Christensen J, Langberg H. The MIPAM trial: a 12-week intervention with motivational interviewing and physical activity monitoring to enhance the daily amount of physical activity in community-dwelling older adults – a study protocol for a randomized controlled trial. *BMC Geriatr*. 2020;20(1):412. doi:10.1186/s12877-020-01815-1

30. Bandura A. *Self-Efficacy: The Exercise of Control*. W H Freeman/Times Books/ Henry Holt & Co; 1997.

31. Luszczynska A, Schwarzer R. Social cognitive theory. In: *Faculty of Health Sciences Publications*. ; 2015:225-251. https://researchbank.acu.edu.au/fhs_pub/5211

32. Moyers TB, Rowell LN, Manuel JK, Ernst D, Houck JM. The Motivational Interviewing Treatment Integrity Code (MITI 4): Rationale, preliminary reliability and validity. *J Subst Abuse Treat*. 2016;65:36-42. doi:10.1016/j.jsat.2016.01.001

33. Larsen RT, Korfitsen CB, Juhl CB, Andersen HB, Langberg H, Christensen J. Criterion validity for step counting in four consumer-grade physical activity monitors among older adults with and without rollators. *Eur Rev Aging Phys Act*. 2020;17(1):1. doi:10.1186/s11556-019-0235-0

34. Craig CL, Marshall AL, Sjöström M, et al. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35(8):1381-1395. doi:10.1249/01.MSS.0000078924.61453.FB

35. IPAQ scoring protocol - International Physical Activity Questionnaire. Accessed March 6, 2019. https://sites.google.com/site/theipaq/scoring-protocol

36. Macfarlane DJ, Lam TH, Stewart SM. Validity of the International Physical Activity Questionnaire Short Form (IPAQ-SF): a systematic review. *Int J Behav Nutr Phys Act*. 2011;8:115. doi:10.1186/1479-5868-8-115

37. Mai KS, Sandbaek A, Borch-Johnsen K, Lauritzen T. Are lifestyle changes achieved after participation in a screening programme for Type 2 diabetes? The ADDITION Study, Denmark. *Diabet Med J Br Diabet Assoc*. 2007;24(10):1121-1128. doi:10.1111/j.1464-5491.2007.02238.x

38. Danquah IH, Petersen CB, Skov SS, Tolstrup JS. Validation of the NPAQ-short – a brief questionnaire to monitor physical activity and compliance with the WHO recommendations. *BMC Public Health*. 2018;18(1):601. doi:10.1186/s12889-018-5538-y

39. World Health Organization. *Global Recommendations on Physical Activity for Health*. WHO; 2010.

40. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy Amst Neth*. 1990;16(3):199-208.

41. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727-1736. doi:10.1007/s11136-011-9903-x

42. Haywood KL, Garratt AM, Fitzpatrick R. Quality of life in older people: a structured review of generic self-assessed health instruments. *Qual Life Res Int J Qual Life Asp Treat Care Rehabil*. 2005;14(7):1651-1668.

43. Janssen MF, Pickard AS, Golicki D, et al. Measurement properties of the EQ-5D-5L compared to the EQ-5D-3L across eight patient groups: a multi-country study. *Qual Life Res*. 2013;22(7):1717-1727. doi:10.1007/s11136-012-0322-4

44. Russell DW. UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *J Pers Assess*. 1996;66(1):20-40. doi:10.1207/s15327752jpa6601_2

45. Lee J, Cagle JG. Validating the 11-Item Revised University of California Los Angeles Scale to Assess Loneliness Among Older Adults: An Evaluation of Factor Structure and Other Measurement Properties. *Am J Geriatr Psychiatry*. 2017;25(11):1173-1183. doi:10.1016/j.jagp.2017.06.004

46. Resnick B, Jenkins LS. Testing the Reliability and Validity of the Self-Efficacy for Exercise Scale. *Nurs Res*. 2000;49(3).

47. Resnick B. Reliability and validity of the Outcome Expectations for Exercise Scale-2. *J Aging Phys Act*. 2005;13(4):382-394.

48. Due P, Holstein B, Lund R, Modvig J, Avlund K. Social relations: network, support and relational strain. *Soc Sci Med 1982*. 1999;48(5):661-673.

49. Lund R, Nielsen LS, Henriksen PW, Schmidt L, Avlund K, Christensen U. Content validity and reliability of the Copenhagen social relations questionnaire. *J Aging Health*. 2014;26(1):128-150. doi:10.1177/0898264313510033

50. U.S. Food and Drug Administration. What is a Serious Adverse Event? *FDA*. Published online 2019. Accessed July 3, 2020. https://www.fda.gov/safety/reporting-serious-problems-fda/what-serious-adverse-event

51. R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing; 2018. https://www.R-project.org

52. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011;64(4):395-400. doi:10.1016/j.jclinepi.2010.09.012

53. Landi F, Cesari M, Onder G, Lattanzio F, Gravina EM, Bernabei R. Physical activity and mortality in frail, community-living elderly patients. *J Gerontol A Biol Sci Med Sci*. 2004;59(8):833-837. doi:10.1093/gerona/59.8.m833

54. Theou O, Blodgett JM, Godin J, Rockwood K. Association between sedentary time and mortality across levels of frailty. *Can Med Assoc J*. 2017;189(33):E1056-E1064. doi:10.1503/cmaj.161034

55. Tudor-Locke C, Craig CL, Aoyagi Y, et al. How many steps/day are enough? For older adults and special populations. *Int J Behav Nutr Phys Act*. 2011;8:80. doi:10.1186/1479-5868-8-80

56. Lee I-M, Shiroma EJ, Kamada M, Bassett DR, Matthews CE, Buring JE. Association of Step Volume and Intensity With All-Cause Mortality in Older Women. *JAMA Intern Med*. 2019;179(8):1105-1112. doi:10.1001/jamainternmed.2019.0899

57. Sylvia LG, Bernstein EE, Hubbard JL, Keating L, Anderson EJ. A Practical Guide to Measuring Physical Activity. *J Acad Nutr Diet*. 2014;114(2):199-208. doi:10.1016/j.jand.2013.09.018

58. Barrett S, Begg S, O'Halloran P, Kingsley M. Integrated motivational interviewing and cognitive behaviour therapy can increase physical activity and improve health of adult ambulatory care patients in a regional hospital: the Healthy4U randomised controlled trial. *BMC Public Health*. 2018;18(1):1166. doi:10.1186/s12889-018-6064-7

59. Holland AE, Mahal A, Hill CJ, et al. Home-based rehabilitation for COPD using minimal resources: a randomised, controlled equivalence trial. *Thorax*. 2017;72(1):57-65. doi:10.1136/thoraxjnl-2016-208514

60. Koolhaas CM, van Rooij FJA, Schoufour JD, et al. Objective Measures of Activity in the Elderly: Distribution and Associations With Demographic and Health Factors. *J Am Med Dir Assoc*. 2017;18(10):838-847. doi:10.1016/j.jamda.2017.04.017

61. Doherty A, Jackson D, Hammerla N, et al. Large Scale Population Assessment of Physical Activity Using Wrist Worn Accelerometers: The UK Biobank Study. *PloS One*. 2017;12(2):e0169649. doi:10.1371/journal.pone.0169649

62. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *J Clin Epidemiol*. 2014;67(3):267-277. doi:10.1016/j.jclinepi.2013.08.015

63. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Published 2011. Accessed April 11, 2016. http://handbook.cochrane.org/

64. Chen Y, Holahan C, Li X. Loneliness, physical activity, and self-rated health in middle-aged and older adults. *Innov Aging*. 2018;2(Suppl 1):967. doi:10.1093/geroni/igy031.3582

65. Ismail K, Bayley A, Twist K, et al. Reducing weight and increasing physical activity in people at high risk of cardiovascular disease: a randomised controlled trial comparing the effectiveness of enhanced motivational interviewing intervention with usual care. *Heart*. 2020;106(6):447-454. doi:10.1136/heartjnl-2019-315656

66. Savović J, Jones H, Altman D, et al. Influence of reported study design characteristics on intervention effect estimates from randomised controlled trials: combined analysis of meta-epidemiological studies. *Health Technol Assess Winch Engl*. 2012;16(35):1-82. doi:10.3310/hta16350

67. Moustgaard H, Clayton GL, Jones HE, et al. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. *BMJ*. 2020;368. doi:10.1136/bmj.l6802

68. Etnier JL, Karper WB, Park S-Y, Shih C-H, Piepmeier AT, Wideman L. Motivating Mature Adults to be Physically Active. *J Aging Phys Act*. 2017;25(2):325-331. doi:10.1123/japa.2015-0294

69. Gallagher KM. Helping Older Adults Sustain Their Physical Therapy Gains: A Theory-Based Intervention to Promote Adherence to Home Exercise Following Rehabilitation. *J Geriatr Phys Ther 2001*. 2016;39(1):20-29. doi:10.1519/JPT.0000000000000040

70. de Lacy-Vawdon CJ, Klein R, Schwarzman J, et al. Facilitators of Attendance and Adherence to Group-Based Physical Activity for Older Adults: A Literature Synthesis. *J Aging Phys Act*. 2018;26(1):155-167. doi:10.1123/japa.2016-0363

Appendix

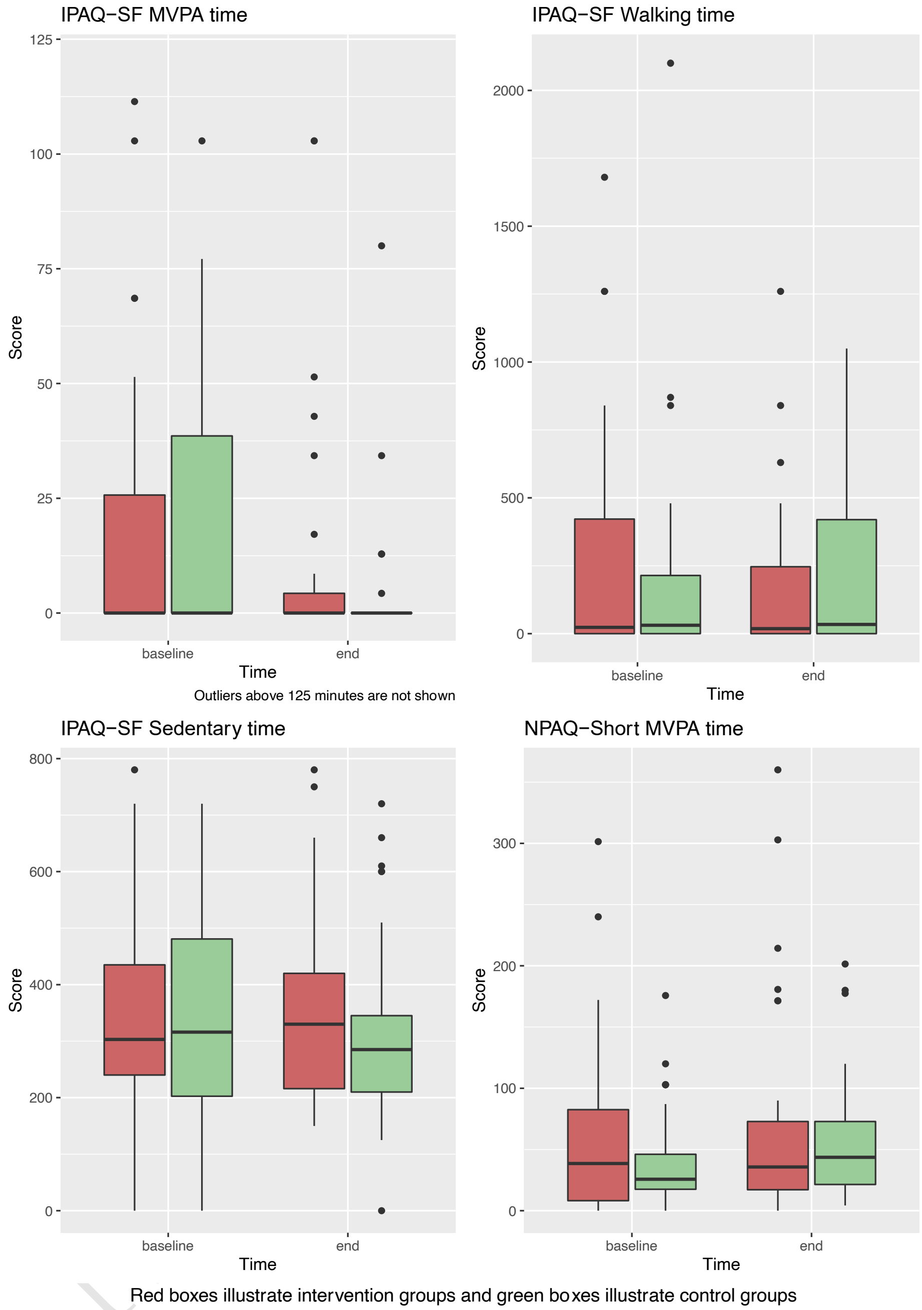


Figure 3. Box plots of secondary self-reported physical activity outcomes at baseline and postintervention endpoint. Red boxes showed the PAM+MI group values and green boxes showed the PAM group values. Thick vertical lines are medians, box size represent interquartile range from 25th to 75th percentile. Notches represent median ± 1.57*IQR/$\sqrt{n}$, and dots represent samples outside this range. Abbreviations: IPAQ-SF: International Physical Activity Questionnaire Short

Form, MVPA: Moderate to vigorous physical activity, NPAQ-Short: Nordic Physical Activity Questionnaire Short. Unimputed data are presented in the boxplot.

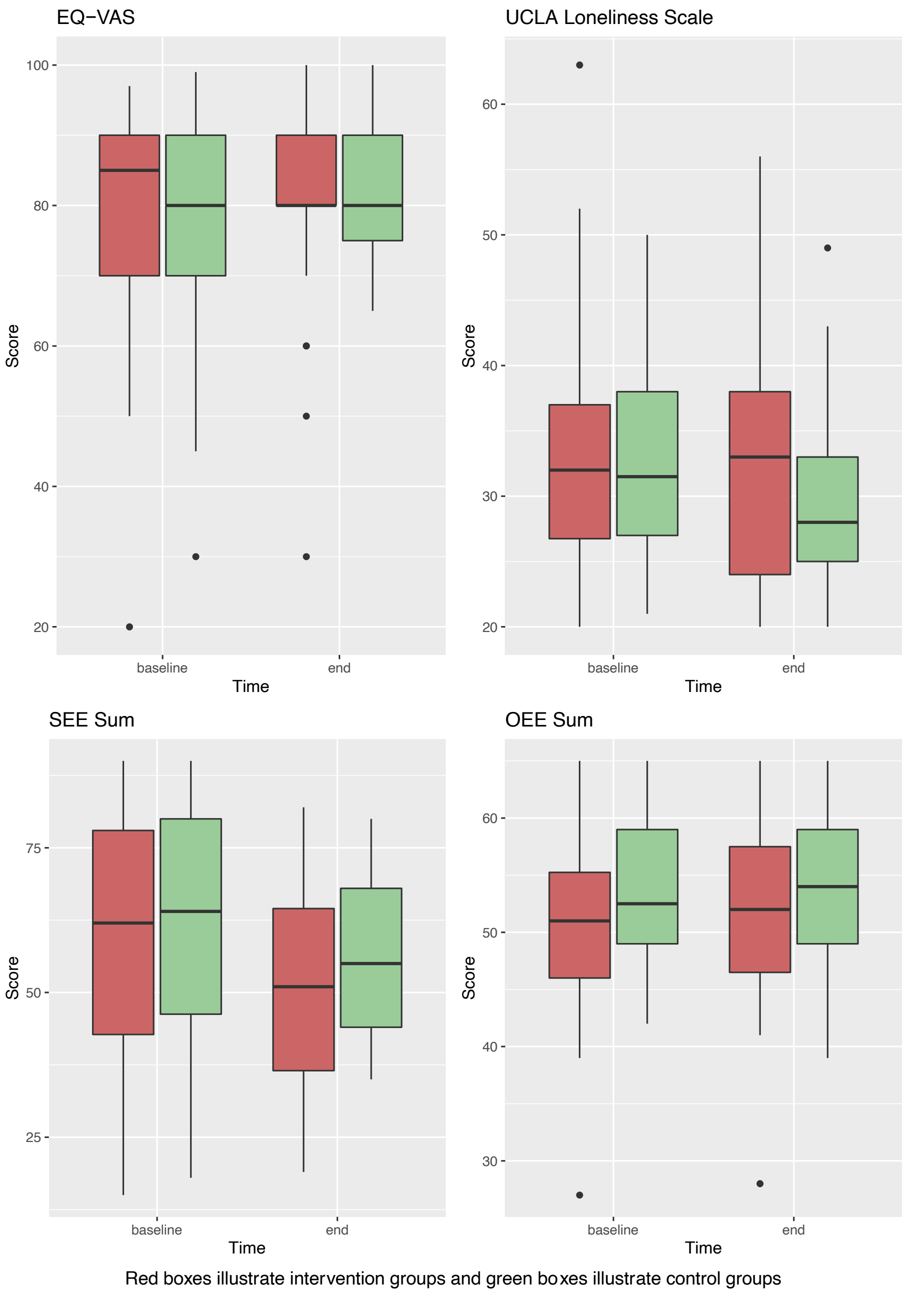


Figure 4. Box plots of secondary self-reported outcomes at baseline and post-intervention. Red boxes showed the PAM+MI group values and green boxes showed the PAM group values. Thick vertical lines are outcome medians, box size represent interquartile range from 25th to 75th

percentile Notches represent median ± 1.57*IQR/$\sqrt{n}$, and dots represent samples outside this range. Abbreviations: EQ-VAS: EuroQol Visual Analogue Scale, UCLA: University of California Los Angeles, SEE: Self-Efficacy for Exercise, OEE: Outcome Expectancy for Exercise-2. Unimputed data are presented in the boxplot.