Self-Report Tools for Assessing Physical Activity in Community-Living Older Adults with Multiple Chronic Conditions: A Systematic Review of Psychometric Properties and Feasibility

Anna Garnett,[1] Jenny Ploeg,[1,2] Maureen Markle-Reid,[1,3] and Patricia H. Strachan[1]

RÉSUMÉ

Identifier l'outil d'autodéclaration de l'activité physique (AP) le plus adapté pour l'évaluation de l'AP chez les personnes âgées vivant dans la communauté qui sont atteintes de multiples maladies chroniques (MMC). L'AP peut avoir une influence positive sur la santé physique et psychologique de cette population. Bien qu'il existe des outils d'auto-évaluation de l'AP, les propriétés psychométriques et la faisabilité de l'utilisation de ces outils chez les personnes âgées avec MMC sont peu connues. Une revue systématique des études publiées entre 2000 et 2018 portant sur les propriétés psychométriques et la faisabilité de 18 outils d'auto-évaluation élaborés pour les personnes âgées vivant en communauté (≥ 65 ans) a été réalisée en vue de déterminer leur pertinence pour les personnes âgées atteintes de MMC. L'évaluation des données disponibles sur les propriétés psychométriques et la faisabilité des 18 outils d'auto-évaluation de l'activité physique a permis d'établir que l'Échelle d'évaluation de l'activité physique chez les personnes âgées (*Physical Activity Assessment Scale for the Elderly*; PASE) est l'outil d'auto-évaluation le plus adapté, qui devrait être recommandé pour la population de personnes âgées avec MMC.

ABSTRACT

The purpose of this study was to identify the self-report physical activity (PA) tool best suited for assessment of PA in community-dwelling older adults with multiple chronic conditions (MCC). PA can positively influence physical and psychological health in this population. Although self-report PA tools exist, little is known about the psychometric properties and feasibility of using these tools in older adults with MCC. A systematic literature review from 2000 to 2018 was conducted of studies reporting on the psychometric properties and feasibility of 18 self-report PA tools for community-dwelling older adults (≥ 65 years) to determine the suitability of these tools for use in older adults with MCC. Based on an assessment of the available evidence for the psychometric properties and feasibility of 18 different self-report PA tools, the Physical Activity Assessment Scale for the Elderly (PASE) is recommended as the best-suited self-report PA tool for older adults with MCC.

La correspondance et les demandes de tirés-à-part doivent être adressées à : / Correspondence and requests for offprints should be sent to:

Anna Garnett, R.N., M.Sc.
Aging, Community, and Health Research Unit
School of Nursing
Faculty of Health Sciences
McMaster University
1280 Main Street West, HSC-3N25
Hamilton, Ontario
Canada L8S 4K1
(garneta@mcmaster.ca)

Introduction

As the world's population is aging, so the prevalence of multiple chronic conditions (MCC) in older adults is also increasing on a global scale (Hwang, Weller, Ireys, & Anderson, 2001; Schellevis, 2013). MCC is commonly defined as having two or more chronic conditions that persist for at least a year (Boyd & Fortin, 2010; Uijen & van de Lisdonk, 2008). For example, in the United States the prevalence of MCC in adults older than 65 years is more than 80 per cent (Buttorff, Ruder, & Bauman, 2017). Older adults with MCC manage conditions such as type 2 diabetes, heart disease, and arthritis, which are associated with poorer health status, higher odds of adverse treatment-related events, and increased health care utilization and associated costs, compared with older adults with single conditions (Bähler, Huber, Brüngger, & Reich, 2015; Canadian Institute for Health Information, 2011; Lehnert et al., 2011; Marengoni et al., 2011; Skinner, Coffey, Jones, Heslin, & Moy, 2016; Tooth, Hockey, Byles, & Dobson, 2008; Wolff, Starfield, & Anderson, 2002). Studies have shown a link between physical inactivity and the development of chronic diseases. The evidence suggests that higher physical activity (PA) may prevent, delay, or even reverse the progression of these diseases and improve the health of community-dwelling older adults with MCC, hereafter referred to as older adults (Clegg, Barber, Young, Iliffe, & Forster, 2014; Miller, Rejeski, Reboussin, Ten Have, & Ettinger, 2000; Public Health Agency of Canada, 2013; Tessier et al., 2000; World Health Organization, 2005).

Enabling older adults to remain in the community is vital to their health, and PA is a critical factor in supporting this goal. Residing at home optimizes older adults' health, autonomy, control, and sense of well-being as well as facilitating their social connectedness (Wiles, Leibing, Guberman, Reeve, & Allen, 2012). A range of PA levels has been shown to increase older adults' mental and physical functioning and it can be instrumental in helping them maintain their cardiovascular, metabolic, and cognitive health as well as their independence (Bouaziz et al., 2017;

McPhee et al., 2016). Furthermore, PA can help reduce or prevent age-related risks such as falls, and limited PA has been associated with care dependency in older adults (Hopewell et al., 2018; Schnitzer et al., 2019). Research suggests that PA levels are reduced in adults with chronic conditions when compared with those of their healthy counterparts (Barker et al., 2019). Therefore, supporting older adults, including those with MCC, to best optimize their health through PA and other health management strategies, is an important goal of health promotion and health management initiatives. Validated PA measures such as self-report PA tools, which can be completed by older adults in their home environment, are well-placed to assess and monitor PA over time. Determining which self-report PA tool is best suited for use in older adults is important to ensure that clinicians are able to accurately assess PA changes in a consistent and reliable manner.

Although many older adults with MCC face challenges related to engaging in PA, PA continues to be a positive influential factor in maintaining or enhancing their physical and psychological health (Bonnefoy et al., 2001). Arthritis diabetes, and heart disease are the most commonly occurring chronic conditions in older adults, and these conditions can be prevented, improved, or modified by PA (Canadian Institute for Health Information, 2011; Tessier et al., 2000; Thompson et al., 2003). Nevertheless, studies have shown that although many older adults with MCC have the ability to meet recommended activity levels for their age, they consistently fail to reach these targets (Ashe, Eng, Miller, & Soon, 2007).

Therefore, increasing older adults' PA levels in support of disease management and health promotion is important for managing chronic conditions in this population (Canadian Institute for Health Information, 2011). Unfortunately, evaluation of PA as part of chronic disease management programs is often hindered by limited availability of reliable data for assessing PA and older adults' PA progress over time (Sun, Norman, & While, 2013).

Background

Direct and objective PA instruments assess PA levels using biological or behavioural markers. The gold standard for directly determining PA is through assessment of total energy expenditure with doubly labeled water (DLW). The use of DLW and other objective PA instruments such as motion sensors, heart rate monitors, and measurement of maximum oxygen capacity (VO_2 max) may be limited by investigator, equipment, participant availability, time constraints, and subject PA capability (Shephard, 2003; Washburn, 2000). Furthermore, the substantial cost and logistics such as delivery and pick-up of PA measurement equipment can restrict their widespread use (Washburn, Smith, Jette, & Janney, 1993). Although other objective measures such as accelerometers and pedometers also provide objective measurements of PA at lower cost, they may not capture PA that involves body parts beyond the lower extremities. In addition, accelerometers also require the availability of advanced computer software to access their data, and pedometers cannot capture changes in behaviour patterns or intensity (Castillo-Retamal & Hinckson, 2011). Validated and reliable self-report PA tools can provide a low-cost, functional means of capturing varying levels of activity over time, making them appropriate for use in community-dwelling older adults with MCC.

Self-report PA tools use recall to provide subjective measures of PA based on the use of questionnaires, diaries, interviews, or other means, reporting on levels and patterns of PA. Self-report PA tools allow for measurement of population-wide activity levels, which can be used to monitor changes in behaviour over time and as a means of testing whether an intervention has resulted in a change in behaviour (Castillo-Retamal & Hinckson, 2011; Prince et al., 2008). Such tools have widespread application in tracking population rates of morbidity and mortality by public health personnel as well as in monitoring progress in the context of an individualized care plan at the level of primary care (Prince et al., 2008). Furthermore they can be used to classify groups of people according to variables such as activity level or duration of exercise, allowing for comparisons among different groups of people and exploration of potential associations between factors such as activity level and insulin sensitivity or triglyceride levels (Lakka et al., 2005; Martins et al., 2017). These indirect measures of PA have increased in popularity because of their affordability, convenience, low participant burden, and ability to capture detailed information about PA, including activities of daily living and leisure-related exercise (Dishman, Washburn, & Shoeller, 2001).

It is increasingly recognized that chronic disease management strongly depends on self-management approaches, of which PA is an essential component (Bodenheimer, Lorig, Holman, & Grumbach, 2002). This means that older adults themselves should also have access to simple, effective, evidence-based PA tools. These tools should be sensitive to a variety of PA activities, including low-intensity PA, which these older adults can use to track their progress, and which are sustainable over time. Pedometers are only able to measure bipedal mobility and cannot capture change in intensity of PA. Accelerometers may also represent challenges when measuring PA in people with mobility issues, because they cannot be used in water and cannot measure activity that does not include acceleration (Kowalski, Rhodes, Naylor, Tuokko, & MacDonald, 2012). In contrast, self-report PA tools are able to assess different types of PA such as swimming or cycling, which are both age-friendly forms of PA. Furthermore, these questionnaires can detect different levels of intensity, and are widely used in research involving population level studies and epidemiological assessments, although there is research suggesting that they may overestimate PA and underestimate inactive time (Schrack et al., 2016; Steene-Johannessen, et al., 2016). Self-report PA tools may be well placed to assess PA in older adults by providing them, as well as their health providers, with an affordable, easily administered alternative that is sensitive to the type of PA engaged in by older adults rather than with costly and complex objective measures such as doubly labelled water or accelerometers. Being able to track PA over time is important in the context of older adults with MCC for various reasons: (1) clinicians need valid and reliable measures of clients' engagement in health management behaviours, (2) monitoring health behaviours can assist older adults in meeting health management goals, and (3) researchers require access to validated, reliable tools in order to better understand the effects of PA; for example, by assessing relative changes in PA at the population level in older adults with MCC (Marengoni et al., 2011; Merom et al., 2014; Portegijs, Sipilä, Viljanen, Rantakokko, & Rantanen, 2017; Vancampfort, Stubbs, & Koyanagi, 2017). However, application of self-report PA tools for all these purposes requires ensuring that these tools are amenable for use and sensitive to the activities engaged in by older adults who may be managing several physical and cognitive challenges (Merom et al., 2014).

Many self-report PA tools are available for use in the adult population (persons ≥ 18 years of age) and several tools have been developed specifically for the older adult population (≥ 65 years of age), such as the Community Healthy Activities Model Program for Seniors (CHAMPS), the Physical Activity Scale for the Elderly (PASE), the Single Self-Report Physical Activity Question (SR-PA), and the Yale Physical Activity

Survey (YPAS) (Bonnefoy et al., 2001; Portegijs et al., 2017; Rosario, Vazques, Cruz, & Ortiz, 2008; Washburn et al., 1993). Research has also shown that PA self-report tools have demonstrated preliminary success as an adjunct to counselling in modifying health behaviours in primary care settings (Petrella & Lattanzio, 2002; Spink & Wilson, 2010). Although to date most PA self-report tools are implemented in the context of research studies, especially larger scale epidemiological research, there is increasing emphasis being placed on the need for clinicians to prescribe and document regular PA as an integral part of primary care (Rippe, 2018). However, little is known about the psychometric properties and feasibility of using self-report PA tools in the older adult population (Bonnefoy et al., 2001). Moreover, self-report PA tools that have been validated in healthy older adults may be ineffective in capturing the amount of and capacity for PA in older adults with MCC. A review by Forsen et al. (2010) reported that although associations were found between self-report PA tool results and health variables, limited knowledge was available on the reliability and validity of these tools in older adults with MCC, a population that engages in limited PA (Hung, Ross, Boockvar, & Siu, 2011). The feasibility of using self-report PA tools for older adults, a critical issue when considering application of such tools in this population, also remains poorly addressed. The current review addresses this knowledge gap by conducting a comprehensive review of the psychometric properties and feasibility of using PA self-report tools in older adults with MCC. The attained results will help inform treating clinicians and older adults with MCC who are engaged in health management strategies about the utility of these tools.

The Review

Aim

The aim of this study was to address the knowledge gaps described, through identification of the self-report PA tool that would be best suited for the assessment of PA in older adults with MCC. This was achieved through performing a review that reported on: (1) the psychometric properties (reliability and validity) of PA self-report tools applied to older adults and (2) the feasibility of using PA self-report tools in older adults with MCC.

Design

The design of this review is that of a systematic review (Aromataris & Pearson, 2014). After development of a clear statement of research aims, inclusion and exclusion criteria were established. Then, a comprehensive search of published studies was conducted applying explicit search terms. The articles captured were assessed for quality (Fayers & Machin, 2007; Voukelatos et al., 2011). Articles of sufficient quality were selected, and data were extracted and then analyzed. Finally, data about various study aspects were presented and synthesized leading to recommendations.

Search Methods

The searched databases were: Cochrane Library, Web of Science, PubMed/MEDLINE®, Cumulative Index of Nursing and Allied Health Literature (CINAHL), and AgeLine. The literature search focused on the period from January 2000 to April 2018, to maximize inclusion of relevant articles. The search terms included: physical activity, measure, assess, quantify, evaluate, tool, instrument, survey, questionnaire, older, elder, retiree, valid, verify, consistent, reliable, property, feasible and community (terms were combined with Boolean operators and equivalent terms searched using wild cards) (Appendix 1).

The search was limited to English language articles reporting on the validity and/or reliability and/or feasibility of the self-report PA tools and that only included subjects ≥ 65 years of age who resided at home or in an unassisted living retirement home. Sixty-five years was chosen as the cut-off for inclusion in the review because it is a commonly used and accepted benchmark of the beginning of older adulthood (Orimo et al., 2006). Articles were excluded if they reported on people residing in long-term care, assisted living settings, or chronic care facilities, or reported on studies that used diaries or journals rather than validated tools to report PA (Appendix 2).

Search Outcome

One reviewer assessed the abstracts and full-text articles for study inclusion. Reference lists of selected articles were hand searched for additional key sources. A total of 1,702 articles were identified through the database searches, with 1,487 remaining after duplicates ($n = 215$) were removed. A total of 1,402 articles were removed because they did not address PA in community-dwelling older adults or did not evaluate a PA tool or did not include a focus on PA assessment. A total of 85 full-text articles were assessed for study eligibility, with the remaining articles ($n = 66$) removed because they did not meet the age inclusion criteria (≥ 65 years) or were not self-report PA tools. A total of 19 studies reporting on 18 self-report PA tools met the inclusion criteria and were included in the review (Figure 1).

Quality Appraisal

A quality appraisal was conducted by one reviewer on all 19 articles using assessment tools adapted from
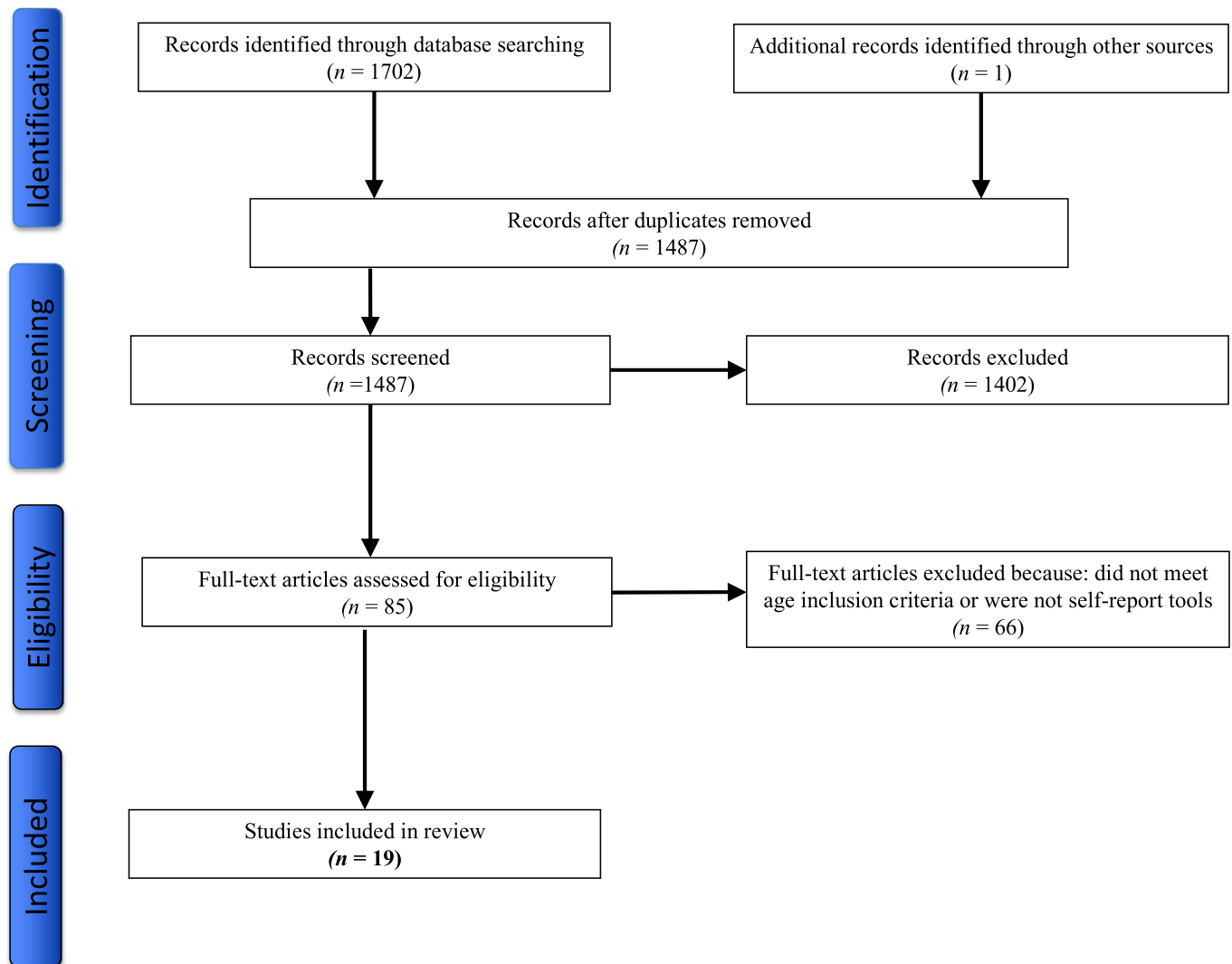
Figure 1: Process of article selection for methodological assessment.

Fayers and Machin (2007) and Voukelatos et al. (2011). All articles were appraised using these criteria, and were deemed of suitable quality and included in this review (Table 1).

Data Abstraction

Psychometric Properties
The 19 selected articles reported on 18 self-report PA tools. The properties of these tools were reported on using the methodological framework by Streiner and Norman (2008) for the following psychometric properties: test–re-test reliability, internal consistency, content validity, construct validity (convergent and discriminatory validity), and criterion validity (concurrent and predictive validity) (Table 2). Although information for all of these properties was sought, it was not available for all self-report PA tools. All tools reported on at least some psychometric properties.

Feasibility
Feasibility is the assessment of whether a plan, activity, or method is practical to implement (Pearson, Field, & Jordan, 2007). For the purpose of this review, feasibility refers to the subjective and objective assessment of the practicality of administering self-report PA tools to older adults with MCC. None of the reviewed articles formally assessed self-report PA tools for feasibility; therefore, the authors conducted an assessment of self-report PA tool feasibility for use in older adults with MCC. Two self-report PA tools were not assessed because the authors were unable to access these tools, although the tool creators and/or recent tool users were contacted. The assessment of PA tool feasibility entailed considering common PA restrictions present in older adults with MCC (e.g., types and intensity of PA), as well as other factors such as the older adults' cognitive status and general wellness, which could affect their ability to complete the self-report tools.

Table 1: Critical appraisal of included articles

Observational Studies (*n* = 15)	Random Sample	Inclusion Criteria	Outcomes & Criteria	Description Comparison Groups	Statistical Analysis	Overall appraisal
Bonnefoy et al., 2001	N	Y	Y	Not specified	Y	Include
Colbert et al., 2011	N	Y	Y	Not specified	Y	Include
Cyarto et al., 2006	N	Y	Y	Not specified	Y	Include
Delbaere, et al., 2010	Y	Y	Y	Not specified	Y	Include
Dinger et al., 2004	N	Y	Y	Not specified	Y	Include
Giles & Marshall, 2009	Y	Y	Y	Not specified	Y	Include
Hagiwara et al., 2008	N	Y	Y	Not specified	Y	Include
Harada et al., 2001	N	Y	Y	Y	Y	Include
Heesch et al., 2011	N	Y	Y	Not specified	Y	Include
Ngai et al., 2012	N	Y	Y	Not specified	Y	Include
Portegijs et al., 2017	Y	Y	Y	Y	Y	Include
Rosario et al., 2008	N	Y	Y	Not specified	Y	Include
Stewart et al., 2001	N	Y	Y	Y	Y	Include
Washburn et al., 1993	Y	Y	Y	Y	Y	Include
Yasunaga et al., 2007	Y	Y	Y	Y	Y	Include

Cohort Studies (*n* = 3)	Represent. Sample	Patients Similar	Outcomes Objective	Sufficient Follow-up	Withdrawals Described	Reliable Measures	Appropriate Statistics	Overall appraisal
Gill et al., 2012	N	Y	U	Y	Y	Y	Y	Include
Godard & Standley 2009	U	Y	Y	Y	Y	Y	Y	Include
Harris et al., 2009	Y	Y	Y	Y	Y	Y	Y	Include

Experimental Studies (*n* = 1)	Random treatment groups	Blinded patients	Blinded allocators[1]	Withdrawals described[1]	Blinded outcome assessors[1]	Comparable treatment & control groups	Identical treatment of groups[1]	Outcomes measured identically[1]	Reliable measures[1]	Sufficient follow-up	Appropriate statistics	Overall appraisal
Merom et al., 2014	Y	N	Y	N	Y	Not specified	Y	Y	Y	N	Y	Include

Table 2: Description of psychometric properties used to assess self-report physical activity tools included in this review

Psychometric Property	Description	Critical Values
Test–retest reliability	Often measured using the correlation coefficient between two sets of responses but can also be measured with Cohen's Kappa (κ) for categorical data.	An intraclass correlation coefficient (ICC) or a κ of 0.7–0.8 is considered good (Fayers & Machin, 2007). Patterson (2000) proposed that time between administrations of self-report physical activity tools should be from 1 to 3 days and not more than 7 days.
Internal consistency	Often measured with the correlation between survey questions intended to measure the same construct.	Cronbach's alpha (α), a common measure of the internal consistency of a tool, should be > 0.7 but not > 0.9 (Streiner & Norman, 2008).
Two types of construct validity	(a) Convergent validity is achieved when two tests intended to measure similar constructs correlate well by testing one tool against the *gold standard* tool with a Pearson's or Spearman's correlation.	The results of the correlation should be high (Trochim, 2005).
	(b) Discriminatory validity exists when two tests that do not measure similar constructs do not strongly correlate (e.g., known-groups validity)	Correlations should be close to 0 (Trochim, 2005).
Two types of criterion validity	(a) Concurrent validity is reached when a new scale correlates with an old scale (often a *gold standard* tool) at the same time.	Both types of validity are measured by correlation with a result closer to 1 being more desirable.
	(b) Predictive validity is the ability of a test to predict something in the future; e.g., an admission test's ability to predict students' performance at university upon graduation.	

The study authors assessed the feasibility of use of the tools with the following criteria developed by Fayers and Machin (2007) and Terwee et al. (2010): (1) Is the tool administered by the respondent or by an interviewer (e.g., health professional)? (2) What is the time required to complete the tool? (3) How many questions are in the tool? (4) Who is the intended population for the tool's application? (5) Are there difficult items in the tool (e.g., additional assistance is necessary to clarify questions, or mistakes are frequently found in completed self-report tools)? (6) Is the tool relevant to the population of interest? (7) Can the tool be completed by a substitute or proxy (e.g., when a respondent is unable to see or write adequately)? (8) What is the recall period of the tool? (9) Does the tool contain integrated instructions for the person completing it? The articles provided some of the information used to assess a tool's feasibility. However, in other cases articles, provided little or no information and the authors of this review accessed the tool directly to enable assessment of tool feasibility.

Synthesis

Many of the articles on the self-report tools reported on multiple different psychometric properties. In order to best collate and present this information, a decision was made to focus on assessments of psychometric properties that pertained to a broad range of PA types, including low-intensity PA. This enabled the widest applicability of the findings to the older adult population, which may engage in limited amounts of PA. Simultaneously, special attention was given to assessment types that occurred across multiple studies and on gold standard measures such as DLW.

The abstracted data on psychometric properties and of feasibility of the 18 self-report PA tools were collated in tables. The available evidence for tools' reliability and validity was reviewed across articles to come to an overall assessment of the tools' psychometric properties and then compared among tools. These comparisons led to the identification of PA tools with the strongest evidence for highest reliability and validity. Subsequently, the available information for tools' feasibility was reviewed and used to identify the tool(s) with highest feasibility. In cases in which the studies provided limited to no information on feasibility, predetermined criteria were used to evaluate the tools' feasibility. In the final step, information on tools' reliability, validity, and feasibility was combined to come to an overall recommendation for the self-report PA tool best suited for application in older adults with MCC.

Results

The 19 articles included in this review provided information on the psychometric properties and feasibility of 18 self-report PA tools (Table 3). Ten of these tools were originally designed for older adults. The remaining eight tools were designed for application to adults (persons ≥ 18 years of age) but applied to older adults. Most of the tools (*n* = 14) were assessed in only one study; one tool, the Incidental and Planned Exercise Questionnaire (IPEQ) was assessed in two studies; another tool, the YPAS, was assessed in three studies;

Table 3: Description of 18 self-report physical activity (PA) tools included in this review

Tool Title & Acronym	Original Purpose	Content & Setting	Dimensions
Active Australia Survey (AAS)	To measure PA in adults 18–75 years of age and to examine knowledge of public health messages about PA health benefits	Walking, house & yard work, recreation, sport	Frequency, duration, intensity
Absolute Physical Activity – Single Question (APA)	To quickly and easily classify absolute and relative PA levels in older adults[a] (minimum age unspecified)	Comprehensive[b]	Level[c]
College Alumni Questionnaire (CAQ)	To measure PA associated with cardiovascular disease in college alumni	Walking, house & yard work, recreation, sport, rest, sleeping	Frequency, duration, intensity
Community Healthy Activities Model Program for Seniors (CHAMPS)	To measure PA levels in underactive older adults ≥ 65 years of age	Walking, house & yard work, recreation, sport, occupational, social	Frequency, duration, intensity
Incidental and Planned Exercise Questionnaire (IPEQ)	To differentiate PA levels between groups of older adults ≥ 65 years of age that differ by fall risk factors	Walking, house & yard work, exercise, sport	Frequency, duration
Lipid Research Clinics Questionnaire (LRCQ)	To assess heavy exertion PA	Work, exercise	Level,[c] frequency
Modified Baecke Questionnaire (MBQ)	To assess habitual PA in community living, older adults 63–80 years of age	Housework, sport, recreation	Frequency, duration, intensity
Modified Dalloso Questionnaire (MDQ)	To assess total PA over a variety of activities	Walking, standing, exercise, leisure	Units[d]
Minnesota Leisure Time Physical Activity Questionnaire (MLTPAQ)	To evaluate energy expenditure in leisure time PA sufficient to condition against cardiovascular disease	House & yard work, recreation, sport	Frequency, duration
Physical Activity Questionnaire – Elderly Japanese (PAQ – EJ)	To explore the typical frequency and duration of PA of elderly Japanese	House & yard work, recreation, sport, transportation	Frequency, duration, intensity
Physical Activity Scale for the Elderly (PASE)	To easily and reliably assess PA of older adults using epidemiologic studies	Recreation, sport, exercise, house & yard work, work	Frequency, duration
Questionnaire d'Activité Physique Saint-Etienne (QAPSE)	To assess habitual PA across a range of activities	Habitual, recreation, sport, exercise, house work, leisure	Duration, intensity
Relative Physical Activity – Single Question (RPA)	To quickly and easily classify absolute and relative PA levels in older adults[1] (minimum age unspecified)	Comprehensive[b]	Level[c]
Seven Day Recall (SDR)	To quantify PA habits of communities, which is practical for large health surveys	Exercise, work, recreation, walking, sleep	Duration, intensity
Single Question Self-report (SR-PA)	To determine whether changes in SR-PA coincide with changes in mobility in 75–90 year olds living independently	House & yard work, recreation, sport, transportation	Duration, intensity
Stanford Usual Activity Questionnaire (SUAQ)	To estimate level of participation in moderate and vigorous PA	Walking, transportation, sport, exercise	Frequency, duration
Yale Physical Activity Survey (YPAS)	To assess a spectrum of low intensity PA in older adults 60–86 years	House & yard work, caretaking, exercise, recreation	Duration
Zutphen Physical Activity Questionnaire (ZPAQ)	To assess PA associated with coronary heart disease risk factors and mortality in Dutch retired men	Walking, exercise, house & yard work, sport, recreation, sleep	Frequency, duration, intensity

Note. [a]Canadian veterans of World War II/the Korean War or their caregivers; [b]Referring to all possible settings; [c]Level of activity, not further specified; [d]Units not further specified.

and two tools, the CHAMPS and the PASE, each were assessed in seven studies (Table 4).

Reliability

Eleven of the 19 articles reported on test–re-test reliability of the tools using four different measures (Table 4): (1) nine articles used intraclass correlation (ICC); (2) two used Cohen's Kappa coefficient (κ); and (3) one used Pearson's correlation coefficient (r). A total of seven tools were assessed for test–re-test reliability, whereas 11 were not (Table 4).

The reported ICC ranged from 0.75 to 0.91 for studies that used a three to seven-day test–re-test time frame and ranged from 0.60 to 0.78 for test–re-test time frames longer than seven days (Table 4). For the shorter test–re-test time frames, the lowest ICC was found for CHAMPS, with a re-test time frame of seven days (ICC = 0.75), whereas the highest ICC was found for PASE, with a test–re-test time frame of three days (ICC = 0.91). For the longer test–re-test time frames (longer than seven days), PASE and CHAMPS were virtually tied for lowest ICC (PASE: ICC = 0.60, test–re-test time frame = 10 days; CHAMPS: ICC = 0.62, test–re-test time frame = 21–49 days). The highest ICC for the longer test–re-test timeframe was found for CHAMPS (ICC = 0.93), with a test–re-test time frame of 7–14 days (Table 4).

Table 4: Psychometric properties of 18 self-report PA tools included in this review as reported by 19 reviewed articles

Source	Study Sample[a]	Tool Acronym[b]	Reliability	Validity		
				Convergent Validity (Spearman's ρ)		
				DLW TEE	**TEE/RMR**	**VO$_2$max**
Bonnefoy et al., 2001	19 men, 73.4 years ± 4.1, Lyon, France	CAQ	NA	0.37 ns	0.42 ns	0.17 ns
		LRCQ	NA	0.29 ns	0.37 ns	0.08 ns
		MBQ	NA	0.28 ns	0.14 ns	0.11 ns
		MDQ	NA	0.23 ns	0.24 ns	0.16 ns
		MLTPAQ	NA	0.17 ns	0.44 ns	0.38 ns
		PASE	NA	0.34 ns	0.32 ns	0.09 ns
		QAPSE	NA	0.25 ns	0.16 ns	0.22 ns
		SDR	NA	0.51*	0.37 ns	0.12 ns
		SUAQ(mod.)	NA	0.46 ns	0.38 ns	0.35 ns
		SUAQ(vig.)	NA	0.64*	0.66*	0.61*
		YPAS	NA	0.10 ns	0.03 ns	0.13 ns
			ICC (10 days)	**Convergent Validity (Spearman's ρ)**		
				DLW/RMR/PAI	**ACC**	
Colbert et al., 2011	56 adults (44 females), 74.7 years ± 6.5, 98% white, Madison, WI, healthy	CHAMPS	0.64§	0.28*	0.52**	
		PASE	0.60§	0.20ns	0.36**	
		YPAS	0.73§	0.07ns	0.37**	
			ICC (1 week) **K (1 week)**	**Predictive Validity (Spearman's ρ)**		
				SFT (4 parts)	**SF-12$_{phys}$**	
Cyarto et al., 2006[c,d]	163 adults (132 females), 79.1 years ± 6.3, Brisbane, Australia, 72% had ≥ 3 chronic conditions (arthritis, hypertension, elevated cholesterol)	CHAMPS	0.75* 0.68**	0.14-0.32**	0.12-0.24**	
			ICC (1 week)	**Concurrent Validity (Spearman's ρ)**		
			IPEQ-W **IPEQ-WA**	**IPEQ-W & IPEQ-WA**		
Delbaere et al., 2010	500 adults (279 females), 77.4 years ± 6, Sydney, Austarlia, healthy, ability to walk unassisted, 28.3% reported falls in past year	IPEQ	0.77§ 0.84§	0.67§-1.00§		
			ICC (3 days)	**Convergent Validity (Spearman's ρ)**		
				PASE & Act		
Dinger et al., 2004	56 adults (43 females), 75.7 years ± 7.9, Caucasian 92.9%, USA	PASE	0.91§	0.43**		
			ICC (1-2 weeks)	**Convergent Validity (Spearman's ρ)**		
				CHAMPS & PED		
Giles & Marshall, 2009[c]	73 adults (45 females), ≥ 65years, Queensland, Australia, healthy, no falls history	CHAMPS	0.57§-0.88§	0.21ns-0.57**		
			K (33-37 days)	**Convergent Validity (Spearman's ρ)**		
			APA **RPA**	**APA & interRAI** **RPA & interRAI**		
Gill et al., 2012	159 adults (females 58), 80 years ± 3.9, ON, 38% reported one or more falls in past year	APA and RPA	0.75§ 0.56§	0.10§-0.33§ 0.28§-0.57§		
				Concurrent Validity (Pearson's r)		
				AAHPERD		
Godard & Standley, 2006	50 adults (41 female), 72.96 years ± 7.66, Macomb, IL	CHAMPS	NA	(-0.11) ns – 0.15ns		

Continued

Table 4: Continued

Source	Study Sample[a]	Tool Acronym[b]	Reliability	Validity		
			ICC (3-4 weeks)	**Concurrent Validity (Spearman's ρ)** JALSPAQ		
Hagiwara et al., 2008[c]	325 adults (191 females), 72.6 years ± 4.9, no ADL or cognitive deficiencies, Chugoku, Japan	PASE	0.65[§]	0.48**		
			ICC (3-7 weeks) CHAMPS	**Convergent Validity (Pearson's r)**		
				CHAMPS & 6 min walk	**YPAS & 6 min walk**	**PASE & 6 min walk**
Harada et al., 2001[c,d]	87 adults (53 females), 79 years ± 6, good cognitive status, 49% hypertension, 39% arthritis, 36% visual deficits, 34% white, 58% Asian American, 8% African American, Los Angeles, CA	CHAMPS, PASE & YPAS	0.62[§]	0.46**	0.58**	0.68**
				Convergent Validity (Pearson's r)		
				ACC	**PED**	
Harris et al., 2009[6]	238 adults (112 females), 73.6 years ± 6.1, 29.1% arthritis, 38% ≥ 2 chronic conditions, UK	ZPAQ	NA	0.35***	0.36***	
				Known Groups Validity		
				Less active[e]	**More active (95% CI)**	**Very active (95% CI)**
				0.0	0.71 (0.43-0.99)	1.09 (0.77-1.42)
				Convergent Validity (Spearman's ρ)		
				PED	**SF-36$_{phys}$**	**ST**
Heesch et al., 2011	53 adults (26 females), 72.6 years ± 5.9, able to follow study protocol, Queensland, Australia	AAS	NA	0.42**	0.39**	-0.25[ns]
				Concurrent Validity (Spearman's ρ) ACC		
Merom et al., 2014	492 adults (354 females), ≥ 65 years, New South Wales, Australia, physically inactive, no medical issues precluding exercise, no cognitive impairments	IPEQ	NA	0.09[ns] – 0.23**		
			ICC[f]	**Concurrent Validity (Spearman's ρ)**		
				SF -36$_{phys}$	**SF -36$_{role\ phys}$**	**10-m walk time**
Ngai et al., 2012	90 adults (54 females), 77.7 years ± 7.7, good cognitive status (MMSE score), China	PASE – C (Chinese)	0.81***	0.58**	0.47**	-0.28**
				Concurrent Validity (Spearman's ρ)		
				ACC	**Low intensity PA**	**Moderate intensity PA**
Portegijs et al., 2017	174 adults, 75-90 years, Finland, able to communicate and willing to participate	SR-PA (Finnish)	NA	0.49***	0.40***	0.49***
			Internal Consistency (Cronbach's α)			
Rosario et al., 2008[d]	53 adults, 65-90 years, chronic conditions included, severe psychiatric conditions excluded, San Juan, PR	CHAMPS – S (Spanish)	0.76[§]	NA		

Continued

Table 4: Continued

Source	Study Sample[a]	Tool Acronym[b]	Reliability	Validity		
				Convergent Validity (Pearson's r)		
				6 min Walk	**SF-36$_{phys}$**	
Stewart et al., 2001[c,d]	249 adults (159 females), 74 years ± 6, 9% minorities, 59% arthritis, 36.5% hypertension, San Francisco, CA	CHAMPS	NA	0.21***	0.30***	
			ICC (3-7 weeks)	**Convergent Validity (Pearson's r)**		
				PH	**HR**	**Balance**
Washburn et al., 1993[c]	222 adults, mean age 73 years, Watertown, MA	PASE	0.75§	-0.34**	-0.13**	0.33**
			Pearson's r (4 weeks)	**Concurrent Validity (Spearman's ρ)** **PAQ – EJ & ACC**		
Yasunga et al., 2007[c]	147 adults (86 females), 72.8 years ± n4.25 years, good cognitive status, Nakanojo, Tokyo	PAQ – EJ (Japanese)	0.70§	0.41***		

Note. ^{ns} $p \geq 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, §p not reported.

[a] Sample size, gender, age, ethnicity, PA ability. [b]For explanation of tool acronyms see Table 2. [c]Study also reviewed in Forsen et al. (2010). [d]Study includes subjects with multiple chronic conditions. [e]Referent group. [f]Time frame not reported.

AAS = Active Australia Survey; AAHPERD = American Alliance for Health, Physical Education, Recreation and Dance Functional Fitness Assessment; ACC = accelerometer; Act =Actigraph Accelerometer; APA = absolute physical activity question; CAQ = College Alumni Questionnaire; CHAMPS = Community Healthy Activities Model Program for Seniors; CI = confidence interval; DLW/RMR/PAI = doubly labeled water/resting metabolic rate/physical activity index; DLW TEE = doubly labeled water total energy expenditure; HR = heart rate; ICC = intraclass correlation coefficient; interRAI = interRAI Community Health Assessment Form; IPEQ = Incidental and Planned Exercise Questionnaire; JALSPAQ = Japan Arteriosclerosis Longitudinal Study Physical Activity Questionnaire; K = Cohen's kappa; LRCQ = Lipid Research Clinics Questionnaire; MBQ = Modified Baecke Questionnaire; MDQ = Modified Dalloso Questionnaire; MLTPAQ = Minnesota Leisure Time Physical Activity Questionnaire; PA = physical activity; PAQ-EJ = Physical Activity Questionnaire, Elderly Japanese; PASE = Physical Activity Assessment Scale for the Elderly; PED = pedometer; PH = perceived health; RPA = relative physical activity question; SDR = Seven Day Recall; SF-12$_{phys}$ = Short-form 12 Health Survey questionnaire, physical activity component; SF-36$_{phys}$ = Short-form 36 physical activity component; SF-36$_{role\ phys}$ = Short-form 36 role physical; SFT = Senior Fitness Test; SR-PA = Single Self-Report Physical Activity Question; ST = step test; SUAQ = Stanford Usual Activity Questionnaire; TEE/RMR = total energy expenditure by resting metabolic rate; VO$_2$max = maximal oxygen uptake; YPAS = Yale Physical Activity Survey.

Results for κ ranged from 0.56 for the Relative Physical Activity (RPA) tool with a test–re-test time frame of 33–37 days, to 0.75 for the Absolute Physical Activity (APA) tool with a test–re-test time frame of 33–37 days (Table 4). Both articles reported Cohen's κ for the whole scale total, and only reported unweighted assessments of κ.

Validity

Eighteen of the 19 articles reported on validity using four different measures (Table 4): (1) ten articles reported on convergent validity; (2) seven reported on concurrent validity; (3) one reported on predictive validity; and (4) one reported on known groups validity. Six articles reported on convergent validity using Spearman's correlation coefficient (ρ), and four articles used Pearson's correlation coefficient (*r*). Six articles reported on concurrent validity using Spearman's correlation coefficient (ρ) and one article used Pearson's correlation

coefficient (*r*). One article, Rosario et al. (2008) did not report on validity.

Convergent validity assessed with Spearman's ρ (Table 4) for step counts ranged from 0.21 to 0.57. Both the highest and lowest values were found for CHAMPS (ρ = 0.21–0.57), with intermediate values for PASE (ρ = 0.36–0.43) and the Active Australia Survey (AAS) (ρ = 0.42). Convergent validity assessed with Spearman's ρ (Table 4) for DLW ranged from 0.07 for the YPAS to 0.64 for the Stanford Usual Activity Questionnaire (SUAQ$_{(vig.)}$). Both PASE (ρ = 0.20–0.34) and CHAMPS (ρ = 0.28) showed intermediate values. Using other questionnaires to validate the tools resulted in correlations (ρ) of 0.39 for the AAS and 0.10–0.57 for the APA/RPA.

Convergent validity assessed with Pearson's *r* (Table 4) for the six-minute walk test ranged from 0.21 to 0.68. The lowest value was found for CHAMPS (*r* = 0.21) and the highest value was found for PASE (*r* = 0.68).

When compared to step counts, Pearson's *r* was 0.35–0.36 for the Zutphen Physical Activity Questionnaire (ZPAQ). Using other questionnaires to validate the tools resulted in correlations (*r*) ranging from 0.30 for CHAMPS to -0.34 for PASE (Table 4).

Concurrent validity assessed with Spearman's ρ (Table 4) for PASE ranged from 0.47 to 0.58 when the tool was compared with other questionnaires, and was -0.28 (reverse scale) when compared with walking speed. For IPEQ and the Physical Activity Questionnaire (PAQ), ρ ranged from 0.09 to 0.23 and 0.41, respectively, when compared with accelerometers. For SR-PA, ρ ranged from 0.40 to 0.49, when compared with accelerometers and time spent on low- or moderate-intensity PA.

Feasibility

Sixteen of the 18 tools were available for access allowing an assessment of feasibility aspects for use in older adults with MCC beyond the aspects reported by the reviewed articles (Table 5). The two remaining self-report PA tools (Modified Dalloso Questionnaire [MDQ] and Questionnaire d'Activité Physique Saint-Etienne [QAPSE]) could not be accessed despite attempts to make contact with the tool creators and/or recent tool users (Bonnefoy et al., 2001). Because this lack of access resulted in very incomplete information, these tools had to be excluded from further feasibility assessment.

Seven of the reviewed self-report PA tools are interviewer-administered, six are self-administered, and three can be either self-administered or interviewer-administered (Table 5). The time required for completing the tools ranged from less than 5 minutes for the APA, Lipid Research Clinics Questionnaire (LRCQ), RPA, and SR-PA to more than 30 minutes for the Minnesota Leisure Time Physical Activity Questionnaire (MLTPAQ) (Bonnefoy et al., 2001; Gill, Jones, Zou, & Speechley, 2012; Portegijs et al., 2017) . The number of questions per tool ranged from 1 (SR-PA) to 63 (MLTPAQ) (Bonnefoy et al., 2001; Portegijs et al., 2017). Eight tools contained some questions that were difficult to answer by study participants as reported by the article authors (Table 5). Two tools (LRCQ and SUAQ) were not relevant to older adults because they did not address activities of daily life (e.g., housework or gardening) that are important in this population (Bonnefoy et al., 2001). Most articles (*n* = 13) did not specify whether the tools could be completed by a substitute for the intended subject, whereas three stated explicitly that a substitute was not permitted (Table 5). Six tools had a recall time frame of one week, whereas the recall time frame of the remaining tools ranged from four weeks to 12 months (Table 5).

Table 5: Feasibility aspects of 18 self-report PA tools included in this review

Tool Acronym[a]	Manner of Administration	Required Time	Number of Questions	Intended Population	Difficult Items	General Relevance[b]	Completion by Substitute[c]	Recall Period	Integrated Instructions
AAS	Self	~10 min	9	18-75 years	No	Yes	No	1 week	Yes
APA	Interview	<5 min	2	≥ 65 years	No	Yes	No	Not specified	Yes
CAQ	Self	~15-20 min	8	≥ 18 years	Yes	Yes	Not specified	Day/wk/yr	Yes
CHAMPS	Self or interview	~15-30 min	40	≥ 65 years	Yes	Yes	Not specified	4 weeks	Yes
IPEQ	Self	5-10 min	10	≥ 65 years	Yes	Yes	Not specified	3 months	Yes
LRCQ	Interview	<5min	4	≥ 18 years	No	No	Not specified	Not specified	Yes
MBQ	Self or interview	~10 min	8	63-80 years	Yes	Yes	Not specified	12 months	Yes
MDQ[d]	NA	NA	NA	NA	NA	NA	NA	NA	NA
MLTPAQ	Interview	>30 min	63	≥ 18 years	Yes	Yes	Not specified	12 months	Yes
PAQ – EJ	Self	~20 min	14	≥ 65 years	No	Yes	Not specified	1 week	Yes
PASE	Self or interview	~5 min	10	≥ 65 years	Yes	Yes	Not specified	1 week	Yes
QAPSE[d]	NA	NA	NA	NA	NA	NA	NA	NA	NA
RPA	Interview	<5 min	2	≥ 65 years	No	Yes	No	Not specified	Yes
SR-PA	Self	<5 min	1	≥ 65 years	No	Yes	Not speechly	6 months	Yes
SDR	Interview	~20 min	15	≥ 18 years	Yes	Yes	Not specified	1 week	Yes
SUAQ	Interview	~10 min	9	≥ 18 years	No	No	Not specified	1 week[e]	Yes
YPAS	Interview	~20-30 min	27	≥ 65 years	No	Yes	Not specified	1 week	Yes
ZPAQ	Self	~15 min	17[f]	≥ 65 years	Yes	Yes	Not specified	1 month	Yes

Note. [a]For explanation of tool acronyms see Table 3 [b]Questions cover all activities of daily life not just formalized PA. [c]Questionnaire can be completed by another person on one's behalf. [d]Unable to gain access to these tools or information about them, therefore feasibility aspects of these tools could not be assessed. [e]Compared with previous 3 months. [f]17 main items but 29 items when counting sub-items.

Discussion

To our knowledge, this is the first systematic review reporting on the reliability, validity, and feasibility of self-report PA tools for community-dwelling older adults with MCC, with the aim of identifying the most suitable tool for use in this population. Overall, 19 articles were reviewed reporting on 18 tools that were assessed for reliability, validity, and feasibility.

Psychometric Properties

Self-report PA tools must undergo rigorous evaluation of their psychometric properties in order to have meaningful application to inform research and clinical practice. These tools must be reliable and valid in their ability to detect PA in the population of interest: in this case, older adults with MCC. However, this review determined that reliability was reported on less than half of the self-report PA tools. All articles reporting ICC as a measure of reliability indicated moderate to excellent reliability for the tools they tested. PASE and CHAMPS were found to have the highest reliability, and both were reported to have excellent or close to excellent reliability (Koo & Li, 2016). However, a range of factors including the exact model used to calculate this statistic affects ICC values (Koo & Li, 2016; Lee et al., 2012). Unfortunately, only one of the reviewed articles provided information about the exact ICC model used (for IPEQ), making comparison of reliability among the various tools more difficult (Delbaere, Hauer, & Lord, 2010).

Information on convergent validity was provided for 15 of the self-report PA tools. The results suggest that the six tools showing highest convergent validity (≥ 0.50 for: CHAMPS, PASE, RPA, Seven Day Recall [SDR], SUAQ, YPAS) strongly overlap in their validity values. CHAMPS, PASE and YPAS were assessed most often for convergent validity, providing greater evidence for their validity, but showing a wide range of values for all three tools (Table 4). Convergent validity for all three tools tended to be higher when the tools were compared with direct behavioural measures (e.g. an accelerometer) than compared with direct metabolic measures (e.g., DLW) (Colbert, Matthews, Havighurst, Kim, & Schoeller, 2011; Harada, Chiu, King, & Stewart, 2001; Ngai, Cheung, Lam, Chiu, & Fung, 2012).

Information on concurrent validity was provided for five self-report PA tools. Concurrent validity for PASE and IPEQ was reported by two articles, but was reported by only one article for the SR-PA, Physical Activity Questionnaire – Elderly Japanese (PAQ-EJ) and CHAMPS, thus making available more evidence for validity assessment of PASE and IPEQ. By far the highest values for concurrent validity were found for

IPEQ (Delbaere et al., 2010). However, these values were found by comparing two versions of the IPEQ with each other, instead of with another independent tool. When comparing the IPEQ with PA measured with accelerometers, concurrent validity was much lower (Merom et al., 2014). Excluding the previously mentioned concurrent validity values for IPEQ, the highest and most consistent validity values were found for PASE (Hagiwara, Ito, Sawai, & Kazuma, 2008; Ngai et al., 2012). Overall, the evidence for moderate to high concurrent validity was strongest for PASE.

Particular areas of concern regarding self-report PA tools relate to their ability to detect change in health status over time. The current review found few studies that assessed the tools' sensitivity to change over time, and those that did used variable time frames and comparison modalities, making it difficult to compare and contrast across studies (Godard & Standley, 2006; Harada et al., 2001; Portegijs et al., 2017; Stewart et al., 2001). Although designed with older adults in mind, CHAMPS was shown to be less sensitive to lighter PA levels, limiting its suitability for older adults with MCC (Godard & Standley, 2006).

Feasibility in Older Adults with MCC

The feasibility assessment revealed that a number of self-report PA tools (e.g., PASE and ZPAQ) stood out for generally more favourable feasibility aspects, such as taking less time to administer and requiring a shorter recall period. These tools were deemed by us to be more feasible for use with older adults with MCC than the other assessed tools.

Only five of the reviewed articles explicitly stated that they included older adults with MCC (but excluding cognitive conditions) in their studies (Cyarto, Marshall, Dickinson, & Brown, 2006; Gill et al., 2012; Harris et al., 2009; Portegijs et al., 2017; Stewart et al., 2001). Some articles specified that subjects had to be sufficiently physically healthy to engage in PA such as walking, and all articles stated that subjects had to be free of major cognitive conditions that would affect their ability to self-report PA (Colbert et al., 2011; Giles & Marshall, 2009; Merom et al., 2014;). Although it can be assumed that the ability to respond appropriately to self-report PA tools is limited by major cognitive impairment, it is less clear whether this is also always the case for mild cognitive impairment. Everyday function in people with mild cognitive decline may be comparable to that of a control group, although everyday memory recall may be reduced and become worse with increased recall period (Anstey, Eramudugolla, Chopra, Price, & Wood, 2017; Irish, Lawlor, Coen & O'Mara, 2011). Because the prevalence of mild cognitive impairment in the older adult population may be as high as 42 per

cent, the use of self-report PA tools under conditions of mild cognitive impairment requires further study (Tricco et al., 2012).

In addition to limitations in engaging in PA (Harada et al., 2001), older adults with MCC may face physical limitations such as arthritis and declining visual acuity (diabetes related or age related) (Anderson & Horvath, 2004; Canadian Institute for Health Information, 2011; Harada et al., 2001), which could influence their ability to complete a self-report PA tool. One study using PASE reported that a small proportion of subjects needed additional explanations and assistance to complete the questionnaire (Hagiwara et al., 2008). Cyarto et al. (2006) determined that 25 per cent of participants required assistance with CHAMPS, whereas a mail survey study by Giles and Marshall (2009) found that only 60 per cent of returned CHAMPS questionnaires were fully completed. Factors such as enabling completion of the self-report PA tool by a support person, simple wording, and training older adults in the use of the tools may help improve the feasibility of using PA self-report tools in the older adult population.

The length of the recall period is an important aspect of self-report PA tools. Shephard (2003) suggests that recall diminishes with the length of the recall period. However, some of the reviewed self-report PA tools required recall of up to 12 months, with the College Alumni Questionnaire (CAQ) requiring recall of daily, weekly and yearly activities in increments of hours and minutes (Bonnefoy et al., 2001). This recall demand would be challenging for people of all ages, but perhaps particularly so for older adults with MCC who may not engage in regular, regimented schedules of activity.

From a feasibility perspective, two tools, the ZPAQ and PASE stood out for having high feasibility (Bonnefoy et al, 2001; Colbert et al., 2011; Dinger, Oman, Taylor, Vesely, & Able, 2004; Hagiwara et al., 2008; Harada et al., 2001; Harris et al., 2009; Ngai et al., 2012; Washburn et al., 1993). The ZPAQ may be a very feasible tool overall for older adults with MCC. It is self-administered, contains 17 questions (29 when including sub-items), is expected to take 15 minutes to complete, was designed to capture the PA of older adults 65 years or age or more, and has a recall period of one month. However, PASE may have even higher feasibility as it can be self- or interviewer-administered, contains just 10 questions, takes only five minutes to complete, was designed for use in older adults 65 years of age or more, and has a recall period of only one week. Although the CHAMPS self-report tool was also specifically developed for older adults, the fact that it contains 40 items and has a recall period of one month may make it less feasible for older adults with MCC

(Colbert et al., 2011; Cyarto et al., 2006; Giles & Marshall, 2009; Godard & Standley, 2006; Harada et al., 2001; Rosario et al., 2008; Stewart et al., 2001).

Self-Report PA Tool Recommendation

Following evaluation, some self-report PA tools emerged as most suitable for older adults with MCC. The SUAQ was highlighted by Bonnefoy et al. (2001) because it correlated well with DLW. However, the participants in that study were healthy males who were capable of engaging in maximal physical exertion and had no serious illnesses. Because the SUAQ was designed to report moderate to vigorous PA, it may not be surprising that it correlates well with DLW (Bonnefoy et al., 2001). However, whether the SUAQ is suitable to capture the low-intensity PA typical of older adults with MCC is unclear (Bonnefoy et al. 2001). Although PASE does not correlate well with direct measures of energy expenditure such as DLW, it does correlate well with accelerometers and walking ability (Bonnefoy et al., 2001; Colbert et al., 2011; Dinger et al., 2004; Hagiwara, et al., 2008; Harada et al., 2001; Ngai et al., 2012; Washburn et al., 1993). This makes it well suited for older adults with health constraints who are most likely to engage in low-intensity PA. PASE was also determined to be sensitive to change after a six-week physician counselling intervention, making it well placed for use as part of a health management plan aimed at assisting older adults with MCC in managing their health (Marcus et al., 1997).

Both PASE and CHAMPS demonstrated better psychometric properties than other self-report PA tools included in this review. Both tools were shown to have good reliability by at least some studies (Bonnefoy et al., 2001; Colbert et al., 2011; Cyarto et al., 2006; Dinger et al., 2004; Giles & Marshall, 2009; Godard & Standley, 2006; Hagiwara et al., 2008; Harada et al., 2001; Ngai et al., 2012; Rosario et al., 2008; Stewart et al., 2001; Washburn et al., 1993). However, PASE displayed better concurrent validity than CHAMPS (Godard & Standley, 2006; Hagiwara et al., 2008; Ngai et al., 2012). In most cases, PASE also displayed convergent validity comparable to or better than CHAMPS (Bonnefoy et al., 2001; Colbert et al., 2011; Dinger et al., 2004; Giles & Marshall, 2009; Harada et al., 2001; Stewart et al., 2001; Washburn et al., 1993). Additionally, PASE rated better than CHAMPS in terms of feasibility, with just 10 questions and a one-week recall period compared with the 40 questions and a four-week recall period for CHAMPS (Bonnefoy et al., 2001; Colbert et al., 2011; Dinger et al., 2004; Hagiwara et al., 2008; Harada et al., 2001; Ngai et al., 2012; Washburn et al., 1993). Although ZPAQ and PASE are rated similarly high in feasibility, there is not enough evidence available for comprehensive

assessment of the ZPAQ's psychometric properties (Bonnefoy et al., 2001; Colbert et al., 2011; Dinger et al., 2004; Hagiwara et al., 2008; Harada et al., 2001; Harris et al., 2009; Ngai et al., 2012; Washburn et al., 1993). ZPAQ has also been criticized for not containing any questions pertaining to household activities, which are an important source of PA in older adults (Bonnefoy et al., 2001; Colbert et al., 2011; Dinger et al., 2004; Hagiwara et al., 2008; Harada et al., 2001; Harris et al., 2009; Ngai et al., 2012; Washburn et al., 1993). Based on the results of this review, the main recommendation is that PASE is the most appropriate self-report tool to assess PA in older adults with MCC. This recommendation is congruent with the use of PASE as the PA data collection tool for the Canadian Longitudinal Study on Aging (Canadian Longitudinal Study on Aging, 2015).

Relevance to Health Practice and Research

Increasingly greater numbers of older adults with MCC are living independently. Evidence shows that independence in older age benefits their social networks and overall health (Wiles et al., 2012). Moreover, PA has demonstrated success in helping alleviate the risk of falls or further health decline associated with aging (Schnitzer et al., 2019). Health professionals, such as nurses, are directly involved in supporting older adults with MCC in managing their health using strategies such as ongoing health monitoring and administration of care plans that increasingly include regular exercise (Rippe, 2018). In this context, tools for reporting of PA are needed with acceptable levels of reliability and validity that are feasible for use in older adults with MCC. Use of such tools should enable health professionals to monitor their clients' progress and changes in health, and measure whether health outcomes have been achieved (Forsen et al., 2010). Self-report PA tools are inexpensive, easily utilized, and often well suited to assessing a broad range of PA (Dinger et al., 2004; Helmerhorst, Brage, Warren, Besson, & Ekelund, 2012). Consequently, simple, easily administered, and reliable self-report PA tools such as PASE have an important role to play in supporting positive health behaviours and managing chronic conditions in older adults with MCC. These tools also have relevant applications in research such as assessing relative changes in PA levels over time at the population level and in making comparisons among populations (Portegijs et al., 2017).

Strengths and Limitations

Strengths of this review include its extensive coverage of the literature through a comprehensive search of the literature extending from 2000 through 2018. Multiple databases were used to ensure broad representation of the available literature. Conducting an assessment of self-report PA tool feasibility in community-dwelling older adults with MCC, in addition to assessing their validity and reliability, makes this review particularly relevant for informing clinical practice and research.

Limitations of the review were the inclusion of only English articles, which may have limited the diversity and global representativeness of included articles. Further, one reviewer conducted the literature search, data extraction, and evaluation, whereas in ideal circumstances, there would be at least two reviewers. The exclusion of two articles from the feasibility assessment of the PA tools because of an inability to obtain the tools is also a shortcoming of the review.

Overall, articles included in this review were lacking in study populations that were characteristic of older adults with MCC. Often samples consisted of convenience samples with unequal gender balances or only healthy males (Bonnefoy et al., 2001; Colbert et al., 2011; Cyarto et al., 2006; Dinger et al., 2004; Gill et al., 2012; Harada et al., 2001; Merom et al., 2014). Going forward, research is required that includes study samples that are more representative in gender and health, with direct attention given to the systematic study of feasibility and sensitivity to clinically relevant changes in health status in older adults with MCC. Furthermore, many of the tools were not comprehensively assessed for their psychometric properties, requiring further study of self-report PA tools to close this knowledge gap.

Other questions that require further study include asking under which conditions self-report PA tools might be used successfully by older adults with mild cognitive impairment, what the effects are of various chronic conditions on older adults' ability to complete self-report PA, and what is the most suitable recall period in older adults for the use of self-report PA tools.

Conclusion

Eighteen self-report PA tools were reviewed for their psychometric properties and feasibility for use in assessing PA in older adults with MCC. All tools for which reliability was reported ($n = 7$) showed good to excellent reliability in at least some tests. Although reported validity ratings were variable, assessed self-report PA tools ($n = 18$) consistently compared more favourably with other questionnaires and direct behavioural measures (i.e., an accelerometer) than with direct metabolic measures. Although some studies explicitly included older adults with MCC in their study populations, there remains a dearth of research available on the psychometric properties

and feasibility of self-report PA tools for assessing PA in older adults with MCC. Further research that includes older adults with MCC is warranted to understand the utility of these tools for this population. When feasibility was accounted for, the available evidence suggests that PASE is the most suitable self-report PA tool to assess PA in older adults with MCC.

References

Anderson, G., & Horvath, J. (2004). The growing burden of chronic disease in America. *Public Health Reports*, *119*(3), 263–270. doi: 10.1016/j.phr.2004.04.005

Anstey, K. J., Eramudugolla, R. E., Chopra, S., Price, J., & Wood, J. M. (2017). Assessment of driving safety in older adults with mild cognitive impairment. *Journal of Alzheimer's Disease*, *57*(4), 1197–1205. doi: 10.3233/JAD-161209

Aromataris, E., & Pearson, A. (2014). The systematic review: An overview—Synthesizing research evidence to inform nursing practice. *American Journal of Nursing*, *114*(3), 53–58. doi: 10.1097/01.NAJ.0000444496.24228.2c

Ashe, M. C., Eng, J. J., Miller, W. C., & Soon, J. A. (2007). Disparity between physical capacity and participation in seniors with chronic disease. *Medicine and Science in Sports and Exercise*, *39*(7), 1139–1146. doi: 10.1249/mss.0b013e31804d2417

Bähler, C., Huber, C. A., Brüngger, B., & Reich, O. (2015). Multimorbidity, health care utilization and costs in an elderly community-dwelling population: A claims data based observational study. *BMC Health Services Research*, *15*(23), 1–12. doi: 10.1186/s12913-015-0698-2

Barker, J., Smith Byrne, K., Doherty, A., Foster, C., Rahimi, K., Ramakrishnan, R., et al. (2019). Physical activity of UK adults with chronic disease: Cross-sectional analysis of accelerometer-measured physical activity in 96 706 UK Biobank participants. *International Journal of Epidemiology*. doi: 10.1093/ije/dyz148

Bodenheimer, T., Lorig, K., Holman, H., & Grumbach, K. (2002). Patient self-management of chronic disease in primary care. *Journal of the American Medical Association*, *288*(19), 2469–2475. doi: 10.1001/jama.288.19.2469

Bonnefoy, M., Normand, S., Pachiaudi, C., Lacour, J. R., Laville, M., & Kostka, T. (2001). Simultaneous validation of ten physical activity questionnaires in older men: A doubly labeled water study. *Journal of the American Geriatrics Society*, *49*(1), 28–35. doi:10.1046/j.1532-5415.2001.49006.x

Bouaziz, W., Vogel, T., Schmitt, E., Kaltenbach, G., Geny, B., & Lang, P. O. (2017). Health benefits of aerobic training programs in adults aged 70 and over: A systematic review. *Archives of Gerontology & Geriatrics*, *69*, 110–127. doi: 10.1016/j.archger.2016.10.012

Boyd, C. M., & Fortin, M. (2010). Future of multimorbidity research: How should understanding of multimorbidity inform health system design? *Public Health Reviews*, *32*(2), 451–474. doi: 10.1007/BF03391611

Buttorff, C., Ruder, T., & Bauman, M. (2017). *Multiple chronic conditions in the United States*. Santa Monica, CA: RAND Corporation.

Canadian Institute for Health Information. (2011). Seniors and the Health Care System: What is the impact of multiple chronic conditions? Retrieved from https://secure.cihi.ca/free_products/air-chronic_disease_aib_en.pdf

Canadian Longitudinal Study on Aging. (2015). CLSA Data collection – Maintaining contact questionnaire. Retrieved from https://clsa-elcv.ca/doc/540

Castillo-Retamal, M., & Hinckson, E. A. (2011). Measuring physical activity and sedentary behaviour at work: A review. *Work*, *40*(4), 345–357. doi: 10.3233/WOR-2011-1246

Clegg, A., Barber, S., Young, J., Iliffe, S., & Forster, A. (2014). The home-based older people's exercise trial of a home-based exercise intervention for older people with frailty. *Age and Ageing*, *43*(5), 687–695. doi: 10.1093/ageing/afu033

Colbert, L. H., Matthews, C. E., Havighurst, T. C., Kim, K., & Schoeller, D. A. (2011). Comparative validity of physical activity measures in older adults. *Medicine and Science in Sports and Exercise*, *43*(5), 867–876. doi:10.1249/MSS.0b013e3181fc7162

Cyarto, E. V., Marshall, A. L., Dickinson, R. K., & Brown, W. J. (2006). Measurement properties of the CHAMPS physical activity questionnaire in a sample of older Australians. *Journal of Science and Medicine in Sport*, *9*(4), 319–326. doi:10.1016/j.jsams.2006.03.001

Delbaere, K., Hauer, K., & Lord, S. R. (2010). Evaluation of the incidental and planned exercise questionnaire (IPEQ) for older people. *British Journal of Sports Medicine*, *44*(14), 1029–1034. doi: 10.1136/bjsm.2009.060350

Dinger, M. K., Oman, F., Taylor, E. L., Vesely, S. K., & Able, J. (2004). Stability and convergent validity of the Physical Activity Scale for the Elderly (PASE). *Journal of Sports Medicine and Physical Fitness*, *44*(2), 186–192.

Dishman, R. K., Washburn, R. A., & Schoeller, D. A. (2001). Measurement of physical activity. *Quest*, *53*(3), 295–309.

Fayers, P. M., & Machin, D. (2007). Choosing and scoring questionnaires. In P. M. Fayers & D. Machin (Eds.), *Quality of life* (2nd ed., pp. 209–222). Hoboken, NJ: Wiley.

Forsen, L., Loland, N. W., Vuillemin, A., Chinapaw, M. J .M., Van Poppel, M. N. M., Mokkink, L. B., et al. (2010). Self-administered physical activity questionnaires for the elderly: A systematic review of measurement properties. *Sports Medicine*, *40*(7), 601–623. doi: 10.2165/11531350-000000000-00000

Giles, K., & Marshall, A. L. (2009). Repeatability and accuracy of CHAMPS as a measure of physical activity in a

community sample of older Australian adults. *Journal of Physical Activity & Health*, 6(2), 221–229.

Gill, D. P., Jones, G. R., Zou, G., & Speechley, M. (2012). Using a single question to assess physical activity in older adults: A reliability and validity study. *BMC Medical Research Methodology*, 12(20), 1–10. doi: 10.1186/1471-2288-12-20

Godard, M. P., & Standley, C. M. (2006). Relationship between CHAMPS physical activity questionnaire and functional fitness outcomes in older adults. *Activities, Adaptation and Aging*, 31(1), 19–40. doi: 10.1300/J016v31n01_02

Hagiwara, A., Ito, N., Sawai, K., & Kazuma, K. (2008). Validity and reliability of the physical activity scale for the elderly (PASE) in Japanese elderly people. *Geriatrics & Gerontology International*, 8(3), 143–151. doi:10.1111/j.1447-0594.2008.00463.x

Harada, N. D., Chiu, V., King, A. C., & Stewart, A. L. (2001). An evaluation of three self-report physical activity instruments for older adults. *Medicine and Science in Sports and Exercise*, 33(6), 962–970. doi:10.1097/00005768-200106000-00016

Harris, T. J., Owen, C. G., Victor, C. R., Adams, R., Ekelund, U., & Cook, D. G. (2009). A comparison of questionnaire, accelerometer, and pedometer: Measures in older people. *Medicine and Science in Sports and Exercise*, 41(7), 1392–1402. doi: 10.1249/MSS.0b013e31819b3533

Heesch, K. C., Hill, R. L., van Uffelen, J. G. Z., & Brown, W. J. (2011). Are Active Australia physical activity questions valid for older adults? *Journal of Science and Medicine in Sport*, 14(3), 233–237. doi: 10.1016/j.jsams.2010.11.004

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(103), 1–55. https://doi.org/10.1186/1479-5868-9-103

Hopewell, S., Adedire, O., Copsey, B. J., Boniface, G. J., Sherrington, C., Clemson, L., et al. (2018). Multifactorial and multiple component interventions for preventing falls in older people living in the community. *Cochrane Database of Systematic Reviews*, 7, CD012221. doi: 10.1002/14651858.CD012221.pub2

Hung, W. W., Ross, J. S., Boockvar, K. S., & Siu, A. (2011). Recent trends in chronic disease, impairment and disability among older adults in the United States. *BMC Geriatrics*, 11(47), 1–12. doi: 10.1186/1471-2318-11-47

Hwang, W., Weller, W., Ireys, H., & Anderson, G. (2001). Out-of-pocket medical spending for care of chronic conditions. *Health Affairs*, 20(6), 267–278. doi: 10.1377/hltaff.20.6.267

Irish, M., Lawlor, B. A., Coen, R. F., & O'Mara, S. M. (2011). Everyday episodic memory in amnestic mild cognitive impairment: A preliminary investigation. *BMC Neuroscience*, 12(80), 1–13. doi: 10.1186/1471-2202-12-80

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. doi: 10.1016/j.jcm.2016.02.012

Kowalski, K., Rhodes, R., Naylor, P., Tuokko, H., & MacDonald, S. (2012). Direct and indirect measurement of physical activity in older adults: A systematic review of the literature. *International Journal of Behavioral Nutrition and Physical Activity*, 9(148), 1–21. doi: 10.1186/1479-5868-9-148

Lakka, T. A., Lakka, H., Rankinen, T., Leon, A. S., Rao, D. C., Skinner, J. S., (2005). Effect of exercise training on plasma levels of C-reactive protein in health adults: The HERITAGE Family Study. *European Heart Journal*, 26(19), 2018–2025. doi: 10.1093/eurheartj/ehi394

Lee, K. M., Lee, J., Chung, C. Y., Ahn, S., Sung, K. H., Kim, T. W., et al. (2012) Pitfalls and important issues in testing reliability using intraclass correlation coefficients in orthopaedic research. *Clinics in Orthopedic Surgery*, 4(2), 149–155. doi: 10.4055/cios.2012.4.2.149

Lehnert, T., Heider, D., Leicht, H., Heinrich, S., Corrieri, S., Luppa, M., et al. (2011). Review: Health care utilization and costs of elderly persons with multiple chronic conditions. *Medical Care Research and Review*, 68(4), 387–420. doi: 10.1177/1077558711399580

Marcus, B. H., Goldstein, M. G., Jette, A., Simkin-Silverman, L., Pinto, B. M., Milan, F., et al. (1997). Training physicians to conduct physical activity counseling. *Preventative Medicine*, 26(3), 382–388. https://doi.org/10.1006/pmed.1997.0158

Marengoni, A., Angleman, S., Melis, R., Mangialasche, F., Karp, A., Garmen, A., et al. (2011). Aging with multimorbidity: A systematic review of the literature. *Ageing Research Reviews*, 10(4), 430–439. doi: 10.1016/j.arr.2011.03.003

Martins, J. C., Aguiar, L. T., Nadeau, S., Scianni, A. A., Teixeira-Salmela, L. F., & Coelho de Morais Faria, C. D. (2017). Measurement properties of self-report physical activity assessment tools in stroke: A protocol for a systematic review. *BMJ Open*, 7(2), e012655. doi: 10.1136/bmjopen-2016-012655

McPhee, J., French, D. P., Jackson, D., Nazroo, J., Pendleton, N., & Degens, H. (2016). Physical activity in older age: Perspectives for healthy ageing and frailty. *Biogerontology* 17(3), 567–580. doi: 10.1007/s10522-016-9641-0

Merom, D., Delbaere, K., Cumming, R., Voukelatos, A., Rissel, C., Van Der Ploeg, H. P., et al. (2014). Incidental and planned exercise questionnaire for seniors: Validity and responsiveness. *Medicine and Science in Sports and Exercise*, 46(5), 947–954. doi: 10.1249/MSS.0000000000000196

Miller, M. E., Rejeski, W. J., Reboussin, B. A., Ten Have, T. R., & Ettinger, W. H. (2000). Physical activity, functional limitations, and disability in older adults. *Journal of the American Geriatrics Society*, 48(10), 1264–1272.

Ngai, S. P. C., Cheung, R. T. H., Lam, P. L., Chiu, J. K. W., & Fung, E. Y. H. (2012). Validation and reliability of the physical activity scale for the elderly in Chinese population. *Journal of Rehabilitation Medicine*, *44*(5), 462–465. doi: 10.2340/16501977-0953

Orimo, H., Ito, H., Suzuki, T., Araki, A., Hosoi, T., & Sawabe, M. (2006). Reviewing the definition of "elderly". *Geriatrics and Gerontology International*, *6*(3), 149–158. doi: 10.1111/j.1447-0594.2006.00341.x

Patterson, P. (2000). Reliability, validity, and methodological response to the assessment of physical activity via self-report. *Research Quarterly for Exercise and Sport*, *71*(S2), 15–20.

Pearson, A., Field, J., & Jordan, Z. (2007). *Evidence-based practice in nursing and health care: Assimilating research, experience and expertise*. Oxford: Blackwell Publishing.

Petrella, R. J., & Lattanzio, C. N. (2002). Does counseling help patients get active? Systematic review of the literature. *Canadian Family Physician*, *48*(1), 72–80.

Portegijs, E., Sipilä, S., Viljanen, A., Rantakokko, M., & Rantanen, R. (2017). Validity of a single question to assess habitual physical activity of community-dwelling older people. *Scandinavian Journal of Medicine and Science in Sports*, *27*(11), 1423–1430. doi: 10.111/sms.12782

Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Connor Gorber, S., & 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*(56), 1–24. doi: 10.1186/1479-5868-5-56

Public Health Agency of Canada. (2013). Chronic disease risk factors. Retrieved from http://www.phac-aspc.gc.ca/cd-mc/risk_factors-facteurs_risque-eng.php

Rippe, J. M. (2018). Lifestyle medicine: The health promoting power of daily habits and practices. *American Journal of Lifestyle Medicine*, *12*(6), 499–512. doi: 10.1177/1559827618785554

Rosario, M. G., Vazquez, J. M., Cruz, W. I., & Ortiz, A. (2008). Internal consistency of the CHAMPS physical activity questionnaire for Spanish speaking older adults. *Puerto Rico Health Sciences Journal*, *27*(3), 224–228. doi: 10.1249/01.mss.0000273699.00091.93

Schellevis, F. G. (2013). Epidemiology of multiple chronic conditions: An international perspective. *Journal of Comorbidity*, *3*(2), 36–40. doi: 10.15256/joc.2013.3.25

Schnitzer, S., Blüher, S., Teti, A., Schaeffner, E., Ebert, N., Martus, P., et al. (2019). Risk profiles for care dependency: Cross-sectional findings of a population-based cohort study in Germany. *Journal of Aging and Health*. doi: 10.1177/0898264318822364

Schrack, J. A., Cooper, R., Koster, A., Shiroma, E. J., Murabito, J. M., Rejeski, W. J., et al. (2016). Assessing daily physical activity in older adults: Unraveling the complexity of monitors, measures, and methods. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *71*(8), 1039–1048. doi: 10.1093/gerona/glw026

Shephard, R. J. (2003). Limits to the measurement of habitual physical activity by questionnaires. *British Journal of Sports Medicine*, *37*, 197–206. doi: 10.1136/bjsm.37.3.197

Skinner, H. G., Coffey, R., Jones, J., Heslin, K. C., & Moy, E. (2016). The effects of multiple chronic conditions on hospitalization costs and utilization for ambulatory care sensitive conditions in the United States: A nationally representative cross-sectional study. *BMC Health Services Research*, *16*(77), 1–8. doi: 10.1186/s12913-016-1304-y

Spink, K. S., & Wilson, K. S. (2010). Physician counselling and longer term physical activity. *Journal of Primary Care and Community Health*, *1*(3), 173–177. doi: 10.1177/2150131910380421

Steene-Johannessen, J., Anderssen, S. A., van der Ploeg, H. P., Hendriksen, J. M., Donnelly, A. E., Brage, S., et al. (2016). Are self-report measures able to define individuals as physically active or inactive? *Medicine and Science in Sports and Exercise*, *48*(2), 235–244. doi: 10.1249/MSS.0000000000000760

Stewart, A. L., Mills, K. M., King, A. C., Haskell, W. L., Gillis, D., & Ritter, P. L. (2001). CHAMPS physical activity questionnaire for older adults: Outcomes for interventions. *Medicine and Science in Sports and Exercise*, *33*(7), 1126–1141.

Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales: A practical guide to their development and use* (4th ed.). New York: Oxford University Press Inc.

Sun, F., Norman, I. J., & While, A. E. (2013). Physical activity in older people: A systematic review. *BMC Public Health*, *13*(449), 1–17. doi: 10.1186/1471-2458-13-449

Terwee, C. B., Mokkink, L. B., van Poppel, M. N. M., Chinapaw, M. J. M., van Mechelen, W., & de Vet, H. C. W. (2010). Qualitative attributes and measurement properties of physical activity questionnaires. *Sports Medicine*, *40*(7), 525–537.

Tessier, D., Menard, J., Fulop, T., Ardilouze, J., Roy, M., Dubuc, N., et al. (2000). Effects of aerobic physical exercise in the elderly with type 2 diabetes mellitus. *Archives of Gerontology and Geriatrics*, *31*(2), 121–132.

Thompson, P. D., Buchner, D., Pina, I. L., Balady, G. J., Williams, M. A., Marcus, B. H., et al. (2003). Exercise and physical activity in the prevention and treatment of atherosclerotic cardiovascular disease. A statement from the Council on Clinical Cardiology (Subcommittee on Exercise, Rehabilitation, and Prevention) and the Council on Nutrition, Physical Activity, and Metabolism (Subcommittee on Physical Activity). *Circulation*, *107*, 3109–3116. doi: 10.1161/01.CIR.0000075572.40158.77

Tooth, L., Hockey, R., Byles, J., & Dobson, A. (2008). Weighted multimorbidity indexes predicted mortality, health service use, and health-related quality of life in older

women. *Journal of Clinical Epidemiology*, *61*(2), 151–159. doi: 10.1016/j.jclinepi.2007.05.015

Tricco, A. C., Soobiah, C., Lillie, E., Perrier, L., Chen, M. H., Hemmelgarn, B., et al. (2012). Use of cognitive enhancers for mild cognitive impairment: Protocol for a systematic review and network meta-analysis. *Systematic Reviews*, *1*(25), 1–6. doi: 10.1186/2046-4053-1-25

Trochim, W. M. K. (2005). *Research methods: The concise knowledge base*. Cincinnati, OH: Atomic Dog Publishing.

Uijen, A. A., & van de Lisdonk, E. H. (2008). Multimorbidity in primary care: Prevalence and trend over the last 20 years. *European Journal of General Practice*, *14*(Suppl. 1), 28–32. doi: 10.1080/13814780802436093

Vancampfort, D., Stubbs, B., & Koyanagi, A. (2017). Physical chronic conditions, multimorbidity and sedentary behavior amongst middle-aged and older adults in six low- and middle-income countries. *International Journal of Behavioral Nutrition and Physical Activity*, *14*(147), 1–13. doi: 10.1186/s12966-017-0602-z

Voukelatos, A., Merom, D., Rissel, C., Sherrington, C., Watson, W., & Waller, K. (2011). The effect of walking on falls in older people: The 'Easy Steps to Health' randomized controlled trial study protocol. *BMC Public Health*, *11*(888), 1–9. doi: 10.1186/1471-2458-11-888

Washburn, R. A. (2000). Assessment of physical activity in older adults. *Research Quarterly for Exercise and Sport*, *71*(Suppl. 2), S79–S88. doi: 10.1080/02701367.2000.11082794

Washburn, R. A., Smith, K. W., Jette, A. M., & Janney, C. A. (1993). The physical activity scale for the elderly (PASE): Development and evaluation. *Journal of Clinical Epidemiology*, *46*(2), 153–162.

Wiles, J. L., Leibing, A., Guberman, N., Reeve, J., & Allen, R. E. S. (2012). The meaning of "aging in place" to older people. *The Gerontologist*, *52*(3) 357–366. doi.org/10.1093/geront/gnr098

Wolff, J. L., Starfield, B., & Anderson, G. (2002). Prevalence, expenditures, and complications of multiple chronic conditions in the elderly. *Archives of Internal Medicine*, *162*(20), 2269–2276. doi: 10.1001/archinte.162.20.2269

World Health Organization. (2005). Chronic diseases and their risk factors. Retrieved from http://www.who.int/chp/chronic_disease_report/media/Factsheet1.pdf

Yasunaga, A., Park, H., Watanabe, E., Togo, F., Park, S., Shephard, R. J., et al. (2007). Development and evaluation of the physical activity questionnaire for elderly Japanese: The Nakanojo Study. *Journal of Aging and Physical Activity*, *15*(4), 398–411.

Appendix 1

Search Strategy - MEDLINE®

#1	(TS=(measure* OR assess* OR quantif* OR evaluate*)) AND LANGUAGE: (English) DocType=All document types; Language=All languages;
#2	(TS="physical activity") AND LANGUAGE: (English) DocType=All document types; Language=All languages;
#3	(TS=(tool* OR instrument* OR survey* OR questionnaire*)) AND LANGUAGE: (English) DocType=All document types; Language=All languages;
#4	(TS=(older OR elder* OR retire*)) AND LANGUAGE: (English) DocType=All document types; Language=All languages;
#5	(TS=(valid* OR verif* OR consist* OR reliab* OR propert* OR feasib*)) AND LANGUAGE: (English) DocType=All document types; Language=All languages;
#6	(TS=(community OR "community dwelling" OR "community living")) AND LANGUAGE: (English) DocType=All document types; Language=All languages;
#7	#6 AND #5 AND #4 AND #3 AND #2 AND #1 DocType=All document types; Language=All languages;
#8	(TS=(communit* OR "community dwelling" OR "community living")) AND LANGUAGE: (English) DocType=All document types; Language=All languages;
#9	#8 AND #5 AND #4 AND #3 AND #2 AND #1 DocType=All document types; Language=All languages;

Appendix 2

Inclusion and exclusion criteria for studies included in this review

Inclusion Criteria	Exclusion Criteria
• Study participant age ≥ 65 years • Residing at home or unassisted in a retirement home • Reporting on self-report PA tools	• Reported on studies that used diaries or journals not validated tools to report PA • Study participant age < 65 years • Reporting on PA tools but not including self-report PA tools • No reporting on validity or reliability of self-report PA tools