


## REVIEW

## Open Access



# A systematic literature review of reviews on techniques for physical activity measurement in adults: a DEDIPAC study

Kieran P. Dowd<sup>1</sup>, Robert Szeklicki<sup>2</sup>, Marco Alessandro Minetto<sup>3</sup>, Marie H. Murphy<sup>4</sup>, Angela Polito<sup>5</sup>, Ezio Ghigo<sup>3</sup>, Hidde van der Ploeg<sup>6,7</sup>, Ulf Ekelund<sup>8,9</sup>, Janusz Maciaszek<sup>2</sup>, Rafal Stemplewski<sup>2</sup>, Maciej Tomczak<sup>2</sup> and Alan E. Donnelly<sup>10\*</sup> 

## Abstract

The links between increased participation in Physical Activity (PA) and improvements in health are well established. As this body of evidence has grown, so too has the search for measures of PA with high levels of methodological effectiveness (i.e. validity, reliability and responsiveness to change). The aim of this “review of reviews” was to provide a comprehensive overview of the methodological effectiveness of currently employed measures of PA, to aid researchers in their selection of an appropriate tool. A total of 63 review articles were included in this review, and the original articles cited by these reviews were included in order to extract detailed information on methodological effectiveness. Self-report measures of PA have been most frequently examined for methodological effectiveness, with highly variable findings identified across a broad range of behaviours. The evidence-base for the methodological effectiveness of objective monitors, particularly accelerometers/activity monitors, is increasing, with lower levels of variability observed for validity and reliability when compared to subjective measures. Unfortunately, responsiveness to change across all measures and behaviours remains under-researched, with limited information available.

Other criteria beyond methodological effectiveness often influence tool selection, including cost and feasibility. However, researchers must be aware of the methodological effectiveness of any measure selected for use when examining PA. Although no “perfect” tool for the examination of PA in adults exists, it is suggested that researchers aim to incorporate appropriate objective measures, specific to the behaviours of interests, when examining PA in free-living environments.

**Keywords:** Physical activity, Measurement, Assessment, Accelerometry, Questionnaires, Self-report, Motion sensors, Pedometers, Heart rate monitors, Adults

## Background

Physical inactivity is the fourth leading cause of death worldwide [1]. Despite this, PA levels of adults across developed nations remain low and the promotion of regular participation in PA is a key public health priority [2]. Population level PA surveillance relies upon having tools to accurately measure activity across all population sub-groups. In addition to surveillance, it is essential that valid, reliable and sensitive measures of PA are available to practitioners, researchers and clinicians in

order to examine the effectiveness of interventions and public health initiatives. The accurate measurement of PA in adults has relevance not only for refining our understanding of PA-related disorders [3], but also for defining the dose-response relationship between the volume, duration, intensity and pattern of PA and the associated health benefits.

A number of methods are available for the assessment of PA [4]. When selecting a measurement technique, researchers and practitioners need to consider not only feasibility and practicality of the measure, but also the methodological effectiveness, such as the validity, reliability and sensitivity. Validity refers to the degree to which a test measures what it is intended to measure,

\* Correspondence: [alan.donnelly@ul.ie](mailto:alan.donnelly@ul.ie)

<sup>10</sup>Department of Physical Education and Sport Sciences, Health Research Institute, University of Limerick, Limerick, Ireland

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

and is most often investigated by comparing the observed PA variables determined by the proposed measure with another comparable measure [5]. Criterion validity is when a measure is validated against the 'gold standard' measure. Good agreement between the proposed method and the gold standard provides some assurance that the results are an accurate reflection of PA behaviour. Other frequently examined forms of validity are concurrent validity (when two measures that give a result that is supposed to be equal are compared) and construct validity (when two measures that are in the same construct are compared). Reliability refers to the degree to which a test can produce consistent results on different occasions, when there is no evidence of change, while sensitivity is the ability of the test to detect changes over time [5].

In addition to methodological effectiveness, other factors need to be considered when selecting a method for assessing PA and interpreting the findings derived from these methods. Feasibility often drives the selection of the study measures. Some measures are more feasible than others depending on the setting, number of participants and cost. For example, the use of activity monitors to estimate PA may be less feasible in epidemiological studies where large numbers of individuals are being assessed. Reactivity may mean that the act of measuring PA may change a person's behaviour: for example, being observed for direct observation [6] or wearing an activity monitor may cause the participant to alter their habitual PA behaviour [7]. When using self-report measures, social desirability may result in over-reporting of PA among participants keen to comply with the intervention aims [8]. These factors require careful consideration when selecting methods for assessing PA.

Although methods for the measurement of PA have been extensively examined, reviews to date have focused on specific categories of methods (i.e. self-report questionnaires [9–11], specific techniques i.e. Doubly Labelled Water (DLW) [12], smart phone technology [13], motion sensors and heart rate monitors (HRM) [14], pedometers [15] or a comparison of two or more methods [16]). Some reviews looked exclusively at specific PA behaviours (e.g. walking) [17] or focused solely on validity and/or reliability issues [18–20]. Other reviews have concentrated on methods for assessing PA in population subgroups (e.g. individuals with obesity [21] or older adults) [22–30]. Due to the level of variability in how information on measurement properties has been presented, and due to the wide range of different measures examined in existing reviews, it is extraordinarily difficult for researchers to compare and contrast measures of PA in adult populations.

The purpose of this article is to review existing reviews (a *review of reviews*) that have examined the methodological

effectiveness of measures of PA. To aid in the comparison of measurement properties between different PA measures, original papers referred to within each review article were sourced, and additional analysis of these references was completed to enable better comparison and interpretation of findings. This review of reviews (as it will be referred to for the remainder of this article) is intended to provide a concise summary of PA measurement in adults. This work was completed as a component of the European DEDIPAC (DEterminants of DIet and Physical ACTivity) collaboration.

## Methods

### Literature search and search strategy

A systematic search of the electronic databases PubMed, ISI Web of Science, CINAHL, PsycINFO, SPORTDiscus and EMBASE took place in April 2014. The search strategy was developed by two of the authors from examining existing literature reviews, whereby common terminology utilised by published systematic reviews of specific methodologies or narrative reviews of all methodologies were included [4, 5, 31–35]. The developed search strategy was reviewed and agreed on by all members of the review team. The electronic databases were searched for the terms "Physical Activity" AND "Review OR Meta-Analysis" AND "Self-report" OR "Logs" OR "Diaries" OR "Questionnaire" OR "Recall" OR "Objective" OR "Acceleromet\*" OR "Activity Monitor\*" OR "Motion Sensor\*" OR "Pedom\*" OR "Heart Rate Monitor\*" or "Direct Observation" AND "Valid\*" OR "Reliab\*" OR "Reproducib\*" OR "Sensitiv\*" OR "Responsiv\*". The search terms and criteria were tailored for each specific electronic database to ensure consistency of systematic searching. Only articles that were published in peer reviewed journals in the English language and were included in this review.

### Eligibility for inclusion

Although DLW is suggested as the gold standard measure of energy expenditure [36], it has not been included in the search strategy, as its feasibility for use in population surveillance research is limited due to its high cost and participant invasiveness [34]. Due to similar limitations, indirect calorimetry has also not been included in this search strategy. However, reviews that discuss studies which have examined the validity of PA measures against DLW and indirect calorimetry were included. The term Global Positioning System (GPS) was not included as it was felt that the limitations associated with GPS used alone [37] deemed it an inappropriate measure of PA for population surveillance in its current form.

Review articles that focused solely on the methodological effectiveness of measures of PA in clinical populations and in children/adolescents were not included in this review. Reviews identified in this study which described the methodological effectiveness of measures of

PA in both adult and youths were included, but only the adult data extracted and included.

#### Article selection

A single reviewer screened all article titles, with only articles that were clearly unrelated to the review of reviews removed at this level. Two independent reviewers examined the article abstracts. Results were collated and reported to a third reviewer, who made the final decision in the case of conflicting results. The full texts of included articles were reviewed by two reviewers using the same protocol for handling conflicting results. Reference lists of identified articles were reviewed to ensure that no relevant articles were overlooked. The collated list of accepted reviews was examined by three leading PA measurement experts, who identified key reviews they felt were not included. The full screening protocol was repeated for all supplementary articles identified (Fig. 1).

#### Quality assessment

The methodological quality of the systematic reviews was evaluated using the Assessment of Multiple Systematic Reviews (AMSTAR) quality assessment tool [38]. No similar quality assessment tool exists for narrative reviews. The AMSTAR protocol was applied to each article by two researchers with any conflicting results resolved by a third reviewer.

#### Data extraction

Initially, the full text and the reference list of each review article meeting the inclusion criteria was screened by a single reviewer for all references to methodological effectiveness, and each methods paper was sourced, screened and all relevant data extracted. The extracted data included general information about the article, the specific measure of PA examined and the demographic characteristics, including the sample population age, size and gender.

Finally, all relevant information relating to properties of methodological effectiveness (i.e. reliability, validity and sensitivity) was recorded. This included the key methodological details of the study and all relevant statistics used to examine measures of methodological effectiveness.

#### Data synthesis

Data synthesis was conducted separately for each of the PA measurement methods, including general recommendations of the method and its effectiveness indicators. The results extracted from the methods papers were presented in the following order: Validity data is presented as mean percentage difference (MPD) in modified forest plots. Similar to Prince and colleagues (2008), where possible, the MPD was extracted or

calculated from the original articles as  $\left(\frac{\text{Comparison Measure} - \text{Criterion Measure}}{\text{Criterion Measure}}\right) * 100$  [16]. Data points positioned around the 0 mark suggest high levels of validity compared to the reference measure. Data points positioned to the left of the 0 mark suggest an underestimation of the variable in comparison to the reference measure. Data points positioned to the right of the 0 mark suggest an overestimation of the variable in comparison to the reference measure. The further away from the 0 mark the point is positioned, the greater the under/overestimation. Data points 250% greater than or less than the reference measure were capped at 250%, and are marked with an asterisk. Due to the lack of reporting of variance results, and the use of differing and incompatible measurement units, confidence intervals are not reported.

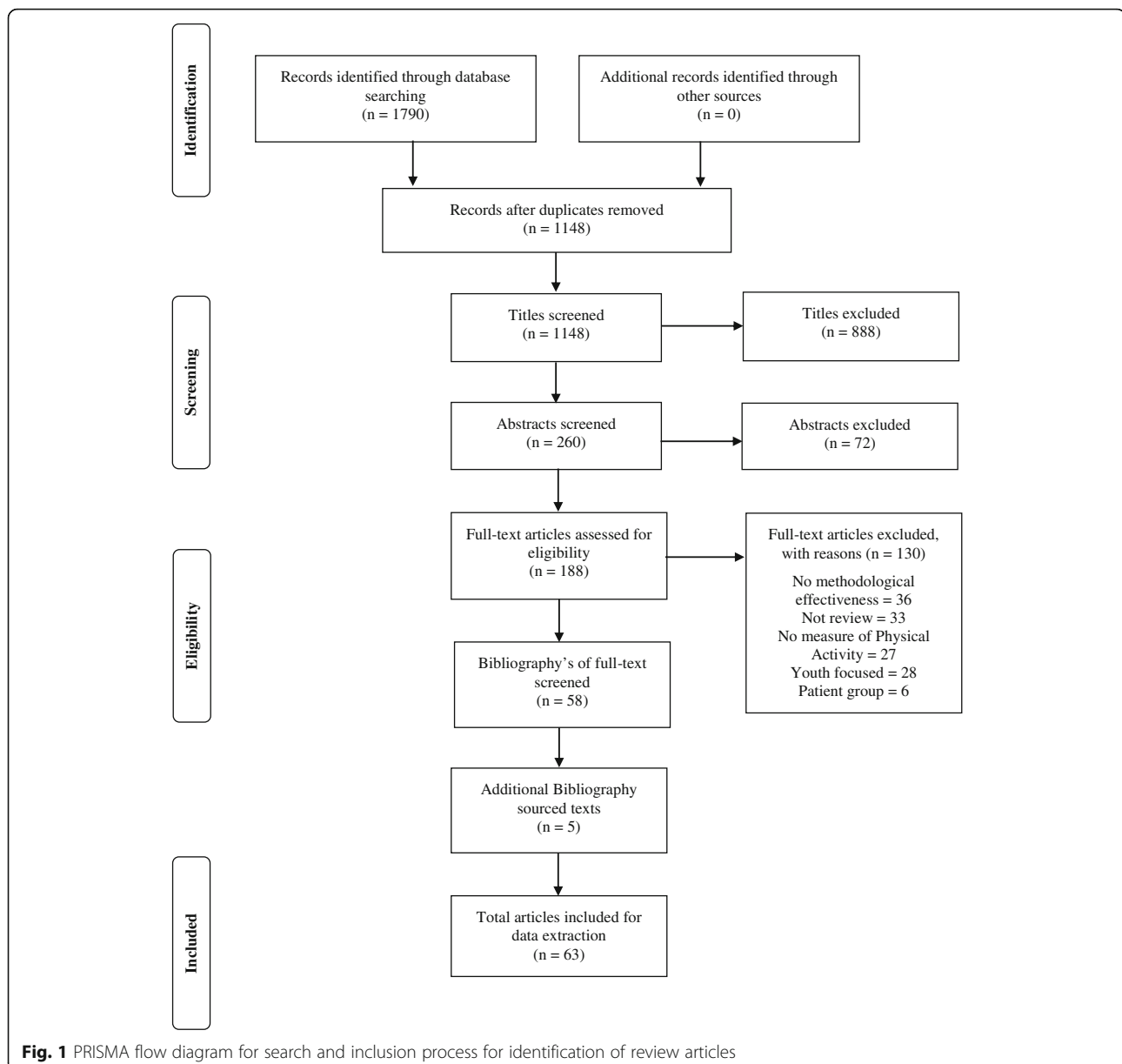
## Results

#### Study selection

The literature search produced 260 potentially relevant abstracts for screening, of which 58 were included in the review following abstract and full text review. After consultation from three international PA experts, and from bibliography review, a further 5 articles were identified for inclusion, providing a total of 63 articles for data extraction (Fig. 1) [4–7, 9–11, 13–19, 21–35, 39–72].

#### Quality assessment

For this article, reviews were categorised as either “Narrative Reviews” or “Systematic Reviews”. A systematic review was defined as a review which described a search strategy for identification of relevant literature. Of the 63 articles, 41 were categorised as narrative reviews, while 22 were identified as systematic reviews. Findings of the AMSTAR quality assessment and review are described in Table 1. The mean AMSTAR score across the 22 articles was 5.4 (out of a possible score of 11), with three articles achieving a score of 3, four articles scoring 4, six articles scoring 5, four articles scoring 6, two articles scoring 7, two articles scoring 8 and one article achieving a score of 9 (Table 1). Based on AMSTAR categorisation, three reviews were considered low quality, 16 reviews were of medium quality and three reviews were considered high quality. The predominant measures examined/discussed in the identified review articles were activity monitors (n=44; 70%), self-report measures (n=28; 44%), pedometers (n=23; 37%) and HRM (n=18; 29%). Other measures included combined physiologic and motion sensors, multi-physiologic measures, multiphasic devices and foot pressure sensors. These measures were incorporated under the combined sensors heading.



## Data extraction

### Self-report measures

#### Validity

**Criterion validity:** A total of 35 articles examined the criterion validity of self-reported measures by comparison to DLW determined energy expenditure [73–107]. Self-reported measures of PA included 7 day recall questionnaires, past year recall questionnaires, typical week questionnaires and PA logs/diaries were validated against 8–15 days of DLW measurement (Additional file 1: Table S1).

The mean values for self-reported and criterion determined PA energy expenditure were available for the calculation of MPD in 27 articles [73–87, 91–93, 95, 97, 99, 100, 102, 104–107]. Energy expenditure was calculated from a

range of behaviours, including leisure time PA, work based PA and PA frequency. The MPD between self-reported PA energy expenditure (time spent in PA normally converted to energy expenditure using a compendium of PA) is presented in Fig. 2. The MPDs observed in studies that examined the validity of PA diaries ranged from -12.9% to 20.8%. MPD for self-reported PA energy expenditure recalled from the previous 7 days (or typical week) were larger, ranging from -59.5% to 62.1%. MPDs from self-reported PA energy expenditure for the previous month compared to DLW determined energy expenditure ranged from -13.3% to 11.4%, while the difference between self-reported PA from the previous twelve months and DLW determined energy expenditure ranged from -77.6% to 112.5%.

**Table 1** Details of the identified reviews, including AMSTAR quality assessment information

Author and Date	Physical Activity Measures of Interest that were Examined	Population Focus	Review Type	AMSTAR
Ainsle et al. (2003) [12]	1. Heart Rate Monitoring 2. Questionnaires and Activity Recalls 3. Pedometers 4. Uniaxial Accelerometers/Activity Monitors 5. Triaxial Accelerometers/Activity Monitors 6. Combined Heart Rate and Motion Sensors	Adult and Youth (age not specified)	Narrative Review	Not Appropriate
Andrew et al. (2010) [72]	1. Questionnaires and Activity Recalls	Adult (age not specified)	Non-Systematic Review	1. Y 2. CA 3. Y 4. N 5. N 6. N 7. N 8. N 9. N 10. N 11. Y Score = 3
Bassett (2000) [20]	1. Heart Rate Monitoring 2. Pedometers 3. Accelerometers/Activity Monitors	Adult and Youth (age not specified)	Narrative Review	Not Appropriate
Bassett et al. (2008) [17]	1. Pedometers 2. Accelerometers/Activity Monitors 3. Direct Observation 4. Questionnaires	Adults (age not specified)	Narrative Review	Not Appropriate
Berlin et al. (2006) [39]	1. Pedometers 2. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Bonomi & Westerterp (2012) [21]	1. Pedometers 2. Accelerometers/Activity Monitors 3. Multi-site Activity Monitors	Adults (age not specified)	Narrative Review	Not Appropriate
Bort-Roig et al. (2014) [13]	1. Smartphone Technology	Adults (age not specified)	Systematic Review	1. Y 2. Y 3. Y 4. N 5. N 6. Y 7. N 8. N 9. N 10. N 11. Y Score = 5
Butte et al. (2012) [40]	1. Pedometers 2. Load Transducers 3. Accelerometers/Activity Monitors 4. Heart Rate Monitors 5. Combined Heart Rate and Motion Sensors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Chen & Bassett (2005) [41]	1. Accelerometers/ Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Cheung et al. (2011) [30]	1. Accelerometers/Activity Monitors	Adults/Older Adults (>17 years)	Systematic Review	1. Y 2. N 3. Y

**Table 1** Details of the identified reviews, including AMSTAR quality assessment information (*Continued*)

Author and Date	Physical Activity Measures of Interest that were Examined	Population Focus	Review Type	AMSTAR
				4. N 5. N 6. Y 7. N 8. N 9. N 10. N 11. Y Score = 4
Corder et al. (2007) [42]	1. Accelerometers/Activity Monitors 2. Pedometers 3. Combined Heart Rate and Motion Sensors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Davidson & deMorton (2007) [43]	1. Self-reported Human Activity Profile	Adults (age not specified)	Systematic Review	1. Y 2. Y 3. Y 4. Y 5. N 6. N 7. N 8. N 9. N 10. N 11. Y Score = 5
DeLany (2012) [44]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Dishman et al. (2001) [6]	1. Direct Observation 2. Questionnaires and Activity Recalls 3. Heart Rate Monitoring 4. Pedometers 5. Accelerometers/Activity Monitors	Adults (age not specified)	Narrative Review	Not Appropriate
Forsen et al. (2010) [29]	1. Self-Administered Physical Activity Questionnaires	Older Adults (mean age > 55 years.)	Systematic Review	1. Y 2. CA 3. Y 4. N 5. N 6. N 7. Y 8. Y 9. CA 10. N 11. Y Score = 5
Freedson & Miller (2000) [14]	1. Pedometers 2. Uniaxial Accelerometers/Activity Monitors 3. Triaxial Accelerometers/Activity Monitors 4. Heart Rate Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Garatachea et al. (2010) [27]	1. Accelerometers/Activity Monitors	Older Adults (age not specified)	Narrative Review	Not Appropriate
Gorman et al. (2014) [28]	1. Accelerometers/Activity Monitors	Older Adults (mean age $\geq$ 65 years or median age >60 years)	Systematic Review	1. Y 2. Y 3. Y 4. N 5. N 6. Y 7. N 8. N

**Table 1** Details of the identified reviews, including AMSTAR quality assessment information (*Continued*)

Author and Date	Physical Activity Measures of Interest that were Examined	Population Focus	Review Type	AMSTAR
				9. Y 11. N 12. Y Score = 6
Haskell et al. (2000) [45]	1. Self-report surveys, Diaries, Physical Activity Logs, Recall Surveys, Retrospective Quantitative History, Global self-report 2. Heart Rate Monitors 3. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Helmerhorst et al. (2012) [31]	1. Physical Activity Questionnaires	Adults and Youth (age not specified)	Systematic Review	1. Y 2. Y 3. Y 4. Y 5. N 6. Y 7. N 8. N 9. NA 10. N 11. Y Score = 6
Kim et al. (2013) [18]	1. International Physical Activity Questionnaire	Adults (15-69 years)	Systematic Review (Meta-analysis)	1. Y 2. Y 3. Y 4. N 5. N 6. Y 7. Y 8. Y 9. Y 10. N 11. Y Score = 8
Kowalski et al. (2012) [32]	1. Self-reported diaries, physical activity logs, physical activity questionnaires, physical activity surveys 2. Accelerometers/Activity Monitors 3. Pedometers 4. Heart Rate Monitoring 5. Direct Observation	Older Adults (mean age > 65 years)	Systematic Review	1. Y 2. Y 3. Y 4. Y 5. N 6. Y 7. Y 8. Y 9. NA 10. N 11. Y Score = 8
Kwak et al. (2011) [19]	1. Physical Activity Questionnaires (assess occupational physical activity)	Adults (age not specified)	Systematic Review	1. Y 2. Y 3. Y 4. N 5. N 6. Y 7. Y 8. Y 9. NA 10. N 11. Y Score = 7
Lamonte & Ainsworth (2001) [46]	1. Accelerometers/Activity Monitors 2. Pedometers	Adults (age not specified)	Narrative Review	Not Appropriate

**Table 1** Details of the identified reviews, including AMSTAR quality assessment information (*Continued*)

Author and Date	Physical Activity Measures of Interest that were Examined	Population Focus	Review Type	AMSTAR
	3. Physical Activity Questionnaires, Records, Logs and Recalls 4. Heart Rate Monitoring			
Lee et al. (2011) [47]	1. International Physical Activity Questionnaire (short form)	Adults and Youth (age not specified)	Systematic Review	1. Y 2. N 3. Y 4. N 5. N 6. Y 7. N 8. N 9. NA 10. N 11. Y Score = 4
Levine (2005) [48]	1. Heart Rate Monitoring 2. Physical Activity Recalls, Logs 3. Pedometers 4. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Liu et al. (2012) [49]	1. Accelerometers/Activity Monitors 2. Foot Pressure Sensors 3. Heart Rate Monitoring 4. Armbands (Consisting of heat flux, galvanic skin response and skin temperature)	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Lowe & O'Laighin (2014) [50]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Mathie et al (2004) [51]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Matthews (2005) [52]	1. Accelerometers/Activity Monitors	Adults (age not specified)	Narrative Review	Not Appropriate
Meyer et al. (2009) [24]	1. Physical Activity Recall Questionnaires, Surveys, Records, Diaries 2. Accelerometers/Activity Monitors 3. Pedometers	Older Adults (mean age >60 years)	Systematic Review	1. Y 2. CA 3. Y 4. N 5. N 6. Y 7. Y 8. Y 9. NA 10. N 11. N Score = 5
Murphy (2009) [25]	1. Accelerometers/Activity Monitors	Older Adults (age not specified)	Narrative Review	Not Appropriate
Neilson et al. (2008) [10]	1. Physical Activity Recall Questionnaires, Surveys, Records.	Adults ( $\geq$ 19 years)	Systematic Review	1. Y 2. CA 3. Y 4. N 5. N 6. Y 7. Y 8. Y 9. NA 10. N 11. N Score = 6
Pedišić et al. (2014) [53]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate



**Table 1** Details of the identified reviews, including AMSTAR quality assessment information (*Continued*)

Author and Date	Physical Activity Measures of Interest that were Examined	Population Focus	Review Type	AMSTAR
Pennathur et al. (2003) [23]	1. Physical Activity Diaries, Questionnaires 2. Accelerometers/Activity Monitors	Older Adults (age not specified)	Narrative Review	Not Appropriate
Pierannunzi et al. (2013) [54]	1. Behavioural Risk Factor Surveillance System	Adults (age not specified)	Systematic Review	1. Y 2. N 3. Y 4. N 5. N 6. N 7. Y 8. N 9. NA 10. N 11. Y Score = 4
Plasqui & Westerterp (2007) [33]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Systematic Review	1. Y 2. CA 3. N 4. N 5. N 6. Y 7. N 8. N 9. NA 10. N 11. Y Score = 3
Plasqui et al. (2013) [55]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Systematic Review	1. Y 2. N 3. N 4. N 5. N 6. Y 7. N 8. N 9. NA 10. N 11. Y Score = 3
Prince et al. (2008) [16]	1. Physical Activity Diaries, Logs, Questionnaires, Surveys and Recall interviews 2. Accelerometers/Activity Monitors 3. Pedometers 4. Heart Rate Monitoring 5. Direct Observation	Adults (mean age > 18 years)	Systematic Review	1. Y 2. Y 3. Y 4. Y 5. N 6. Y 7. Y 8. Y 9. Y 10. N 11. Y Score = 9
Reilly et al. (2008) [56]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Reiser & Schlenk (2009) [34]	1. Physical Activity Diaries, Logs, Questionnaires, Surveys and Recall interviews 2. Direct Observation 3. Pedometers 4. Accelerometers/Activity Monitors 5. Heart Rate Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate

**Table 1** Details of the identified reviews, including AMSTAR quality assessment information (*Continued*)

Author and Date	Physical Activity Measures of Interest that were Examined	Population Focus	Review Type	AMSTAR
Ridgers & Fairclough (2011) [57]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Sallis & Saelens (2000) [9]	1. Physical Activity Diaries, Logs, Questionnaires, Surveys and Recall interviews	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Schutz et al. (2001) [58]	1. Heart Rate Monitors 2. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Shephard (2003) [59]	1. Physical Activity Diaries, Logs, Questionnaires, Surveys and Recall interviews	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Shephard and Aoyagi (2012) [4]	1. Direct Observation 2. Physical Activity Diaries, Logs, Questionnaires, Surveys and Recall interviews 3. Pedometers 4. Uniaxial Accelerometers/Activity Monitors 5. Triaxial Accelerometers/Activity Monitors 6. Mutiphasic Devices 7. Heart Rate Monitoring 8. Multi Physiologic Measures	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Strath et al. (2013) [60]	1. Accelerometers/Activity Monitors	Adults, Older Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Trost et al. (2005) [7]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Tudor-Locke & Myers (2001) [15]	1. Pedometers	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Tudor-Locke & Rowe (2012) [61]	1. Pedometers	Adults (age not specified)	Systematic Review	1. Y 2. CA 3. Y 4. N 5. N 6. Y 7. N 8. N 9. NA 10. N 11. Y Score = 4
Tudor-Locke et al. (2002) [62]	1. Pedometers	Adults and Youth (age not specified)	Systematic Review	1. Y 2. CA 3. Y 4. Y 5. N 6. Y 7. N 8. N 9. NA 10. N 11. Y Score = 5
Tudor-Locke et al. (2004) [63]	1. Pedometers	Adults and Youth (age not specified)	Systematic Review	1. Y 2. CA 3. Y 4. Y 5. N 6. Y 7. N 8. N

**Table 1** Details of the identified reviews, including AMSTAR quality assessment information (*Continued*)

Author and Date	Physical Activity Measures of Interest that were Examined	Population Focus	Review Type	AMSTAR
				9. NA 10. N 11. Y Score = 5
Valanou et al. (2006) [64]	1. Physical Activity Diaries, Logs, Recall Questionnaires, Quantitative History Questionnaires, Global self-report questionnaires 2. Direct Observation 3. Accelerometers/Activity Monitors 4. Pedometers 5. Heart Rate Monitoring	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
van Poppel et al. (2010) [11]	1. Physical Activity Questionnaires	Adults (Mean age 18-55 years)	Systematic Review	1. Y 2. CA 3. Y 4. Y 5. N 6. Y 7. Y 8. Y 9. NA 10. N 11. Y Score = 7
Van Remoortel et al. (2012) [65]	1. Accelerometers/Activity Monitors	Adults (Mean age >18 years)	Systematic Review	1. Y 2. Y 3. Y 4. N 5. N 6. Y 7. N 8. N 9. Y 10. N 11. Y Score = 6
Vanhees et al. (2005) [66]	1. Accelerometers/Activity Monitors 2. Pedometers 3. Heart Rate Monitoring 4. Combined Heart Rate and Movement Sensors 5. Physical Activity Diaries, Logs, Questionnaires, Surveys and Recall interviews	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Warren et al. (2010) [5]	1. Physical Activity Diaries, Logs, Questionnaires, Surveys and Recall interviews 2. Accelerometers/Activity Monitors 3. Heart Rate Monitoring 4. Pedometers	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Washburn (2000) [22]	1. Selected self-reported measures of physical activity	Older Adults (age not specified)	Narrative Review	Not Appropriate
Washburn et al. (2000) [67]	1. Selected self-reported measures of physical activity	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Welk (2005) [68]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Westerterp (2009) [69]	1. Direct Observation	Adults and Youth (age not specified)	Narrative Review	Not Appropriate

**Table 1** Details of the identified reviews, including AMSTAR quality assessment information (*Continued*)

Author and Date	Physical Activity Measures of Interest that were Examined	Population Focus	Review Type	AMSTAR
	2. Physical Activity Diaries, Logs, Questionnaires, Surveys and Recall interviews 3. Heart Rate Monitoring 4. Accelerometers/Activity Monitors			
Westerterp & Plasqui (2004) [70]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate
Yang & Hsu (2010) [71]	1. Accelerometers/Activity Monitors	Adults and Youth (age not specified)	Narrative Review	Not Appropriate

The AMSTAR [38] tool was used to score the reviews as follows: 1. Was an 'a priori' design provided? 2. Was there duplicate study selection and data extraction? 3. Was a comprehensive literature search performed? 4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? 5. Was a list of studies (included and excluded) provided? 6. Were the characteristics of the included studies provided? 7. Was the scientific quality of the included studies assessed and documented? 8. Was the scientific quality of the included studies used appropriately in formulating conclusions? 9. Were the methods used to combine the findings of studies appropriate? 10. Was the likelihood of publication bias assessed? 11. Was conflict of interest reported? Answers: Y=Yes; N=No; CA=Can't answer; NA=Not applicable

**Concurrent validity:** A total of 89 articles reported on concurrent validity of self-reported measures [75, 80, 83, 84, 97, 102, 108–190]. Articles were collated based on the types of referent measures used (Additional file 1: Table S2). The MPD between self-reported energy expenditure and energy expenditure from PA log/diaries for 12 studies ranged from -67.6% to 23.8% (Additional file 1: figure S2a) [80, 108, 110, 111, 128, 145, 152, 157, 159, 160, 169, 175]. These findings suggest that self-report underestimates energy expenditure compared to activity logs/diaries. Seven studies compared self-reported time spent in specific activity intensities with PA intensities from logs/diaries (Additional file 1: figure S2a) [109, 120, 121, 146, 152, 182, 187]. A wider MPD range (-69.0% to 438.5%) was evident, with the greatest differences occurring for moderate intensity physical activity (MPA) and vigorous intensity physical activity (VPA) [109, 120, 121].

Eight studies compared two different self-reported measures of PA energy expenditure [80, 83, 97, 152, 158, 162, 175, 190], and 6 studies compared two different self-reported measures of time spent in PA [112, 135, 136, 146, 152, 153, 158] (Additional file 1: figure S2b). Additional file 1: figure S2c presents 15 studies that compared self-reported PA energy expenditure with PA energy expenditure from activity monitors [80, 132, 142–144, 150, 159, 168, 170, 172, 174, 178, 183, 185, 191]. The MPD ranged from -74.7% to 82.8%, with self-reported measures tending to overestimate energy expenditure.

Self-reported time spent in light intensity physical activity (LIPA) (n=6) [75, 119, 131, 146, 179, 189], MPA (n=17) [75, 115, 119, 130, 131, 133, 134, 139–141, 146, 147, 161, 163, 176, 177, 187, 189] and moderate-to-vigorous intensity physical activity (MVPA) (n=7) [115, 116, 127, 145, 149, 153, 179, 192] was validated against activity monitors that mainly employed count-to-activity thresholds to determine PA intensity (Additional file 1:

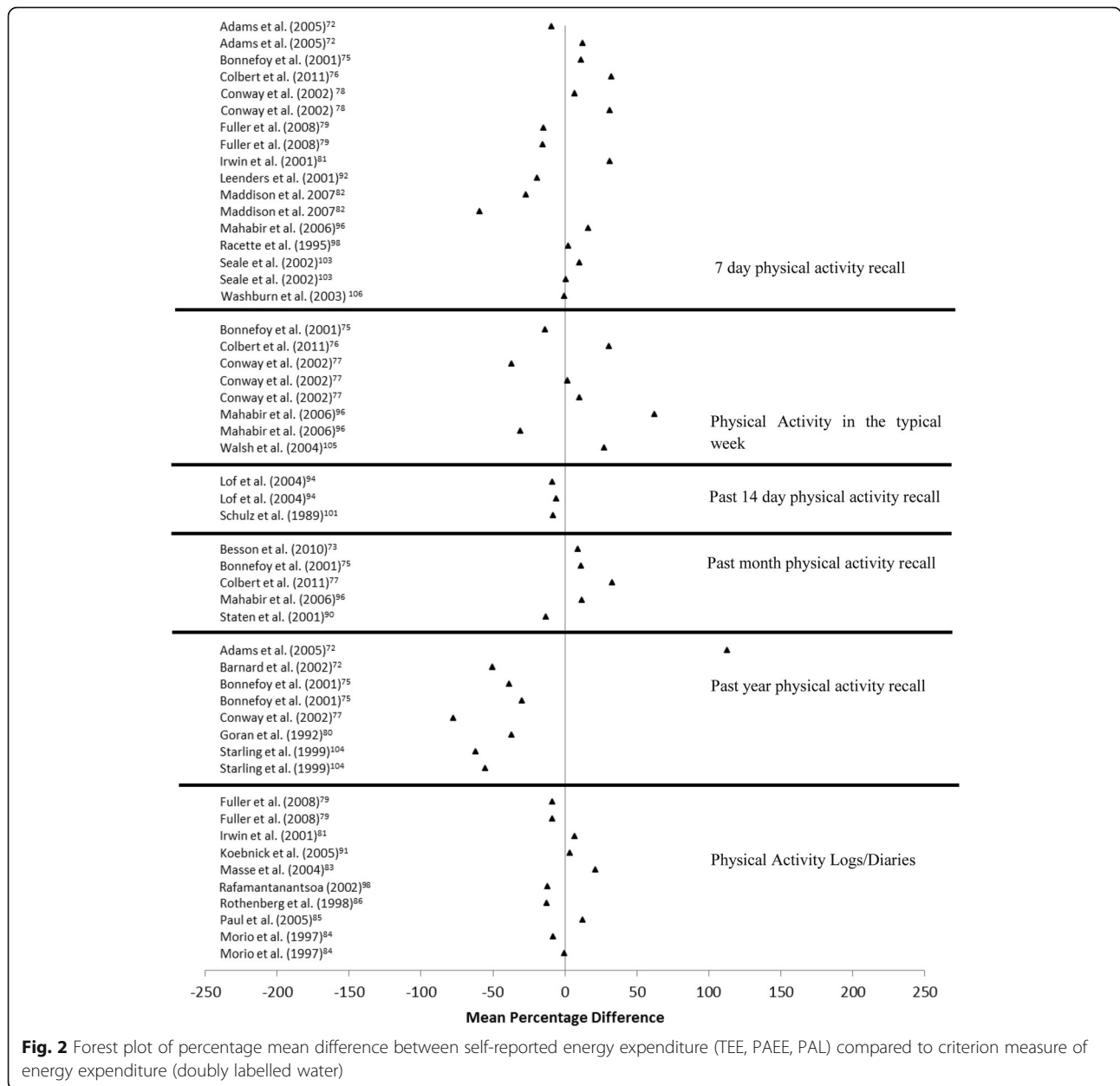
figure S2d), with the MPD for LIPA ranging from -70.1% to 129.2%, MPA ranging from -78.9% to 1007.6% and MVPA ranging from -34.9% to 217.1%. The MPD for VPA was also validated against activity monitors (Additional file 1: figure S2e) [75, 115, 119, 130, 131, 133, 134, 140, 141, 146, 147, 161, 163, 177, 187, 189], with all studies identifying an overestimation of self-reported VPA (Additional file 1: figure S2e).

The concurrent validity of additional self-reported variables, including total PA [163, 181, 184, 193], frequency of MVPA [149], active time [151, 161], time standing [192] and time stepping [192] were also compared to activity monitor determined variables (Additional file 1: figure S2e).

The MPD between self reported energy expenditure and both pedometer and HRM determined energy expenditure [80, 102, 123, 142, 194]; and self-reported time spent in PA intensities and HRM determined time spent in PA intensities [118, 129, 146, 154, 174, 195] are presented in Additional file 1: figure S2f. Self-reported energy expenditure overestimated pedometer determined energy expenditure (range=17.1% to 86.5%). Self-reported measures notably overestimated time spent in PA intensities when compared to HRM. Although self-reported energy expenditure underestimated HRM determined energy expenditure, this underestimation was small compared to other measures (-17.7% to -1.3%).

### Reliability

**Intra-instrument reliability:** One article examined the intra-instrument reliability of a self-reported measure of PA [196]. A self-reported instrument examining the previous 14 days of PA was administered [196]. After 3 days, the instrument examined the PA of the same 14 day period. The findings identified high levels of intra-instrument reliability for total activity (ICC=0.90; 95%



CI=0.86-0.93), MPA (ICC=0.77; 95% CI=0.69-0.84), VPA (ICC=0.90; 95% CI=0.86-0.93), walking (ICC=0.89; 95% CI=0.85-0.93) and energy expenditure (ICC=0.86; 95% CI=0.80-0.90) (Additional file 1: Table S3).

**Test-retest reliability:** The test-retest reliability of self-reported measures was examined in 116 studies [75, 77, 83, 110, 116, 117, 122, 125–127, 129, 131, 132, 135, 137, 140, 144, 145, 147, 149, 152, 153, 155, 157, 159, 161, 162, 167–169, 172, 175, 176, 178–181, 184, 187, 188, 190, 191, 196–269]. Due to the wide test-retest periods, articles were allocated to one of 5 periods, ≤1 week (Additional file 1: Table S4a), >1 - <4 weeks (Additional file 1: Table S4b), >4 - <8 weeks (Additional

file 1: Table S4c), >8 weeks - <1 year (Additional file 1: Table S4d) and >1 year (Additional file 1: Table S4e). Test-retest statistics employed were extracted and are presented in Table 2. An overview of all identified studies examining the test-retest reliability of PA/energy expenditure measured by self-report, along with all test-retest statistics is provided in Additional file 1: Table S4a-e.

**Sensitivity:** Two studies examined the sensitivity of self-reported measures to detect change in PA behaviours over time [256, 270]. Both studies identified small to moderate effect sizes for specific PA behaviours over a six month period in older adults (Additional file 1: Table S5).

**Table 2** Descriptive statistics for reliability of self-reported measures of physical activity across specified time periods

Duration of recall	Correlation Coefficient	Kappa	ICC
≤ 1 week	0.25 – 0.99	0.32 – 0.87	0.30 – 0.99
1 – 4 weeks	0.13 – 0.96	0.40 – 1.00	0.27 – 0.99
5 – 8 weeks	0.41 – 0.99	NA – 0.69	NA
≤ 1 year	0.25 – 0.95	0.54 – 0.82	0.62 – 0.92
> 1 year	0.17 – 0.41	0.20 – 0.85	0.14 – 0.93

ICC Intraclass Correlation Coefficient, NA No data Available

## Activity monitors

### Validity

**Criterion validity:** Fifty-eight articles examined the criterion validity of activity monitor determined PA variables [73, 77, 80, 96, 105, 119, 271–323]. The majority of articles compared activity monitor determined energy expenditure with DLW [73, 77, 80, 96, 105, 274, 275, 277, 278, 280, 281, 285, 292, 293, 295, 296, 300, 303–305, 311, 313, 317, 323], while activity monitor determined steps [119, 271, 283, 287, 289, 298, 299, 306, 307, 314, 315, 318], distance travelled [282] and activity type [272, 273, 276, 279, 284, 286, 288, 290, 291, 294, 297, 301, 302, 308–310, 312, 316, 319–322] were also compared to direct observation (Additional file 1: Table S6).

The range of MPD observed in studies that examined the criterion validity of activity monitor determined energy expenditure ranged from -56.59% to 96.84% (Fig. 3a). However, a trend was apparent for activity monitor determined energy expenditure to underestimate the criterion measure. The range of MPD between activity monitor and direct observation determined steps was smaller, with values ranging from -48.52% to 7.47%, with 96% of studies having a MPD between -10% to 10% (Fig. 3b). Activity monitors overestimated distance walked/run (0.88% to 27.5%). Activity monitors also tended to underestimate activity classification, with MPD varying between -36.67% to 2.00%.

**Concurrent validity:** A total of 103 articles examined the concurrent validity of activity monitor measures of PA [73, 77, 80, 119, 146, 151, 174, 192, 194, 195, 262, 271, 282, 295, 305, 316, 324–409]. Data extractions were grouped by the types of measures used (Additional file 1: Table S7).

The MPD of activity counts from two different activity monitors ranged from -40.6% to 13.2% [262, 327, 351, 389, 392, 405]. The MPD for a wide range of activity behaviours from two different activity monitors were examined; LIPA (-12.5% - 13.7%) [146, 340, 392, 405], MPA (-10.9% - 3.1%) [146, 340], VPA (-9.7% - 20.3%) [146, 352], MVPA (-57.5% - 3.3%) [344, 392, 405], total PA (1.1%) [146]. Stepping [151, 192] and step counts [77, 119, 340, 405] were compared between 2 activity monitor devices (MPD ranged from -21.7% - 0% for step counts and -57.1% - 56% for stepping). Energy

expenditure estimated by two activity monitors were compared [372, 404, 408], with MPD ranging from -21.1% - 61% (Additional file 1: figure S3c).

Energy expenditure at different PA intensities from activity monitors were compared against estimates from indirect calorimetry and whole room calorimetry. For LIPA, the MPD ranged from -79.8% - 429.1% [349, 394]. For MPA, MPD ranged from -50.4% - 454.1% [349, 395], while estimates for VPA ranged from -100% - 163.6%. Energy expenditure estimates from activity monitoring devices for total PA were compared against indirect calorimetry estimates [368, 394, 396, 398, 404], where MPD ranged from -41.4% to 115.7%. The MPD range for activity monitor determined total energy expenditure compared with whole-room calorimetry were narrower (-16.7% to -15.7%) [343, 364] (Additional file 1: figure S3d).

Activity monitor estimates of energy expenditure were compared to HRM estimates of energy expenditure for total PA (-10.4% - 22.2%) [80, 402], for LIPA (-75.4% - 72.8%) [146], for MPA (49.2% - 677.7%), VPA (-46.2% - 46.2%) [146, 361] and for total time spent in PA (-16.1% - 34.9%) [146, 174]. Self-reported measures were used to examine the concurrent validity of activity monitors for energy expenditure [80] and total time spent in PA [174], with MPD ranging from -6.0% - 32.1% (Additional file 1: figure S3e).

Estimated energy expenditure was compared between activity monitors and indirect calorimetry (kcal over specified durations; Additional file 1: figure S3f (-68.5% - 81.1%)) [282, 328, 341, 358, 367, 369, 370, 375, 376, 380, 382, 383, 385, 387]; (METs over specified durations; Additional file 1: figure S3g (-67.3% - 48.4%)) [195, 325, 345–347, 349, 350, 353, 357, 362, 384, 397, 400, 407, 409]. A single study compared the estimated energy expenditure from 5 different activity monitors and indirect calorimetry at incremental speeds (54, 80, 107, 134, 161, 188 and 214 m.min<sup>-1</sup>) in both men and women (MPD ranged from -60.4% - 90.8%) (Additional file 1: figure S3h) [374].

### Reliability

**Inter-instrument reliability:** The inter-instrument reliability of activity monitoring devices (e.g. the reliability of the same device worn by the same participant over the same time period) was examined in 18 studies [301, 315, 337, 344, 370, 385, 387, 406, 409–418]. Study methodologies included the wearing of devices over the left and right hip [337, 370, 385, 387, 406, 413, 415, 417], over the hip and lower back [409], the wearing of devices side by side at the same location on the hip [301, 344, 411, 414, 416–418], devices worn at 3<sup>rd</sup> intercostal space and just below the apex of the sternum [410], device worn on both wrists [412], worn on both legs [315] and worn side by side on the same leg [315]. Coefficients



of variations ranged from 3% to 10.5% for the ActiGraph device [418] and from <6% to 35% for the RT3 accelerometer [387, 416]. All reported correlation coefficients were significant and greater than 0.56 [370, 385, 387, 406, 409, 412, 415, 417]. ICC values for the majority of devices were >0.90 [301, 315, 337, 344, 411, 413], excluding those observed for the RT3 accelerometer (0.72–0.95) [417], Actitrac (0.40–0.87) and Biotrainer devices (0.60–0.71) [406] (Additional file 1: Table S8).

**Test-retest reliability:** Test-retest reliability of activity monitoring devices was examined in 26 studies [153, 155, 228, 234, 262, 282, 297, 314, 358, 385, 407, 411, 414, 416, 419–430]. For the laboratory-based studies, variables examined included distance walked [282], steps at different speeds [314, 420], resting periods [358], accelerometer counts [385, 407, 411, 414, 416, 425, 429, 430], energy expenditure [426] and postural position [297, 429]. For the free-living analyses, behaviours examined included activity behaviours [155, 419], accelerometer counts [262, 421, 422], step count [422], energy expenditure [228, 234] and the number of people achieving the recommended amount of PA [153] (Additional file 1: Table S9).

As the examination of PA over a number of days can be considered a measure of test-retest reliability, researchers have used statistical processes (i.e. generalizability theory or the Spearman Brown Prophecy formula) to determine the minimum number of days required to provide a reliability estimate of PA behaviours [431]. Studies reported that a minimum of three days of ActiGraph data are required to provide a reliable estimate of total PA [423] and time spent in MVPA [424], while a minimum of 2 days is required to provide a reliable estimate of ActiGraph determined steps per day, accelerometer counts per day and intermittent MVPA per day [427]. However, for the examination of continuous 10 minute bouts of MVPA (as suggested in the majority of international PA recommendations), a minimum of 6 days of measurement is required [427].

**Sensitivity:** The only study of responsiveness to change in activity monitors, using the ActiWatch, identified that this device was able to detect significant differences in activity counts accumulated between young adults and sedentary older adults and between active older adults and sedentary older adults [421]. However, no differences could be detected between the young adults and active older adults (Additional file 1: Table S10).

### **Pedometers Validity**

**Criterion validity:** A total of 30 studies were sourced that examined the criterion validity of step count in pedometer devices [283, 289, 298, 306, 307, 314, 318, 365, 391, 432–452], while 3 studies examined the criterion

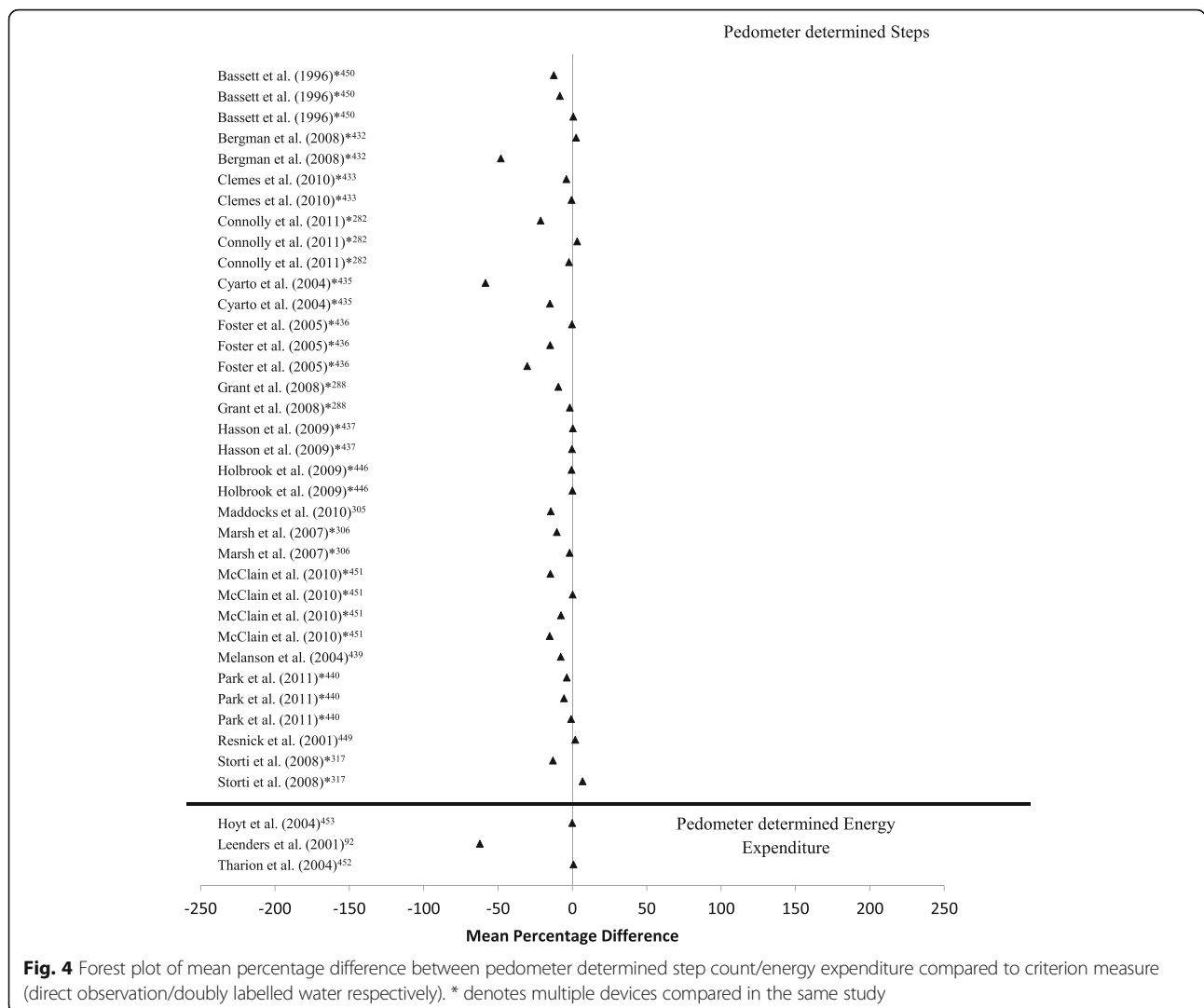
validity of pedometer determined energy expenditure compared to DLW [93, 453, 454]. Of the laboratory based studies assessing criterion validity, 30% used over ground walking protocols [307, 318, 365, 391, 442, 445–447, 450, 451] and the remaining treadmill-based protocols [283, 289, 298, 306, 314, 432–441, 443, 444, 448, 449] or a combination of the two [452]. In free-living studies which examined the criterion validity of pedometer determined energy expenditure, pedometers were worn for 2 [454], 7 [93] and 8 days [453] (Fig. 4; (-62.3% - 0.8%)). Pedometer determined step count was generally lower when compared to direct observation (-58.4% - 6.9%). Some studies also examined the effect of speed on pedometer output. Pedometers had relatively high levels of accuracy across all speeds, but appear to be more accurate at determining step-count at higher walking speeds compared to lower walking speeds (Additional file 1: Table S11) [306, 436, 438, 439].

**Concurrent validity:** The concurrent validity of pedometers was examined in a total of 22 articles [77, 194, 298, 376, 391, 399, 404, 422, 432–434, 441, 444, 448, 449, 451, 452, 455–459]. Various approaches were used to examine the concurrent validity of pedometers, with 14 studies comparing pedometer step count with steps determined from other pedometers [432, 451, 458] and activity monitors [77, 298, 391, 422, 433, 434, 444, 455–457, 459] and 4 studies comparing pedometer determined energy expenditure with energy expenditure determined from indirect calorimetry [376, 399, 404, 441, 448, 451] and/or energy expenditure determined from other activity monitors [451]. One study compared pedometer determined distance travelled with treadmill determined distance travelled [449], while one study compared pedometer determined MVPA with activity monitors determined MVPA [452] (Additional file 1: figure S4a). Pedometers appear to underestimate time spent in MVPA and estimated energy expenditure when compared to other measures. The findings are less clear for step count determined from pedometers when compared to other pedometers or activity monitors, with devices appearing to both over and underestimate step count (Additional file 1: Table S12).

### **Reliability**

**Inter-instrument reliability:** A total of 6 articles examined the inter-instrument reliability of pedometer output obtained from 18 different devices [314, 315, 447, 449, 451, 457]. Many included articles examined the inter-instrument reliability of multiple devices in the same study (e.g. 2 pedometers [315], 5 pedometers [451], 10 pedometers [446, 449]). Inter-instrument reliability was examined by comparing pedometer outputs from two of the same model devices worn on the left and right hip [315, 449, 451, 457], on the left hip, right





hip and middle back [447] and on the left and right hip and repeated with two further devices of the same model [446].

Three studies (1 examining the inter-instrument reliability of a single pedometer and 2 examining the inter-instrument reliability of multiple pedometers), identified that the majority of devices had acceptable levels of inter-instrument reliability ( $ICC \geq 0.80$ ) [446, 449, 457]. In the studies which examined the inter-instrument reliability of multiple devices, 8/10 pedometers [449] and 9/10 pedometers [446] achieved  $ICC \geq 0.80$ . Using planned contrasts, Bassett and colleagues highlight that no significant differences were observed between devices worn on the left and right hip [451]. Two studies investigated the effect of walking speed on inter-instrument reliability, highlighting that  $ICC$  values increased as speed increased [315, 447] (Additional file 1: Table S13).

**Test-retest reliability:** A single laboratory-based test-retest reliability study in a laboratory-based treadmill

protocol identified that the Yamax Digiwalker SW-200 (Tokyo, Japan) had appropriate test-retest reliability ( $ICC > 0.80$  and significant) at 7 out of 11 treadmill speeds (non-significant speeds = 4, 20, 22 and 26  $km \cdot h^{-1}$ ) [314].

A total of 6 articles examined the reliability of pedometer steps obtained over a specified measurement period [423, 427, 460–463], presenting the minimum number of days of pedometer measurement to reliably estimate PA behaviours. The minimum number of days of measurement required for a reliable estimate (i.e.  $ICC > 0.8$ ) of pedometer steps was 2-4 days (Additional file 1: Table S14) [423, 427, 460–463].

**Sensitivity:** In the only study of pedometer responsiveness to change, effect size was used to examine the meaningfulness of difference between means [464]. A large effect size ( $>0.8$ ) was observed, suggesting that pedometers, in this study, were sensitive to change (Additional file 1: Table S15).

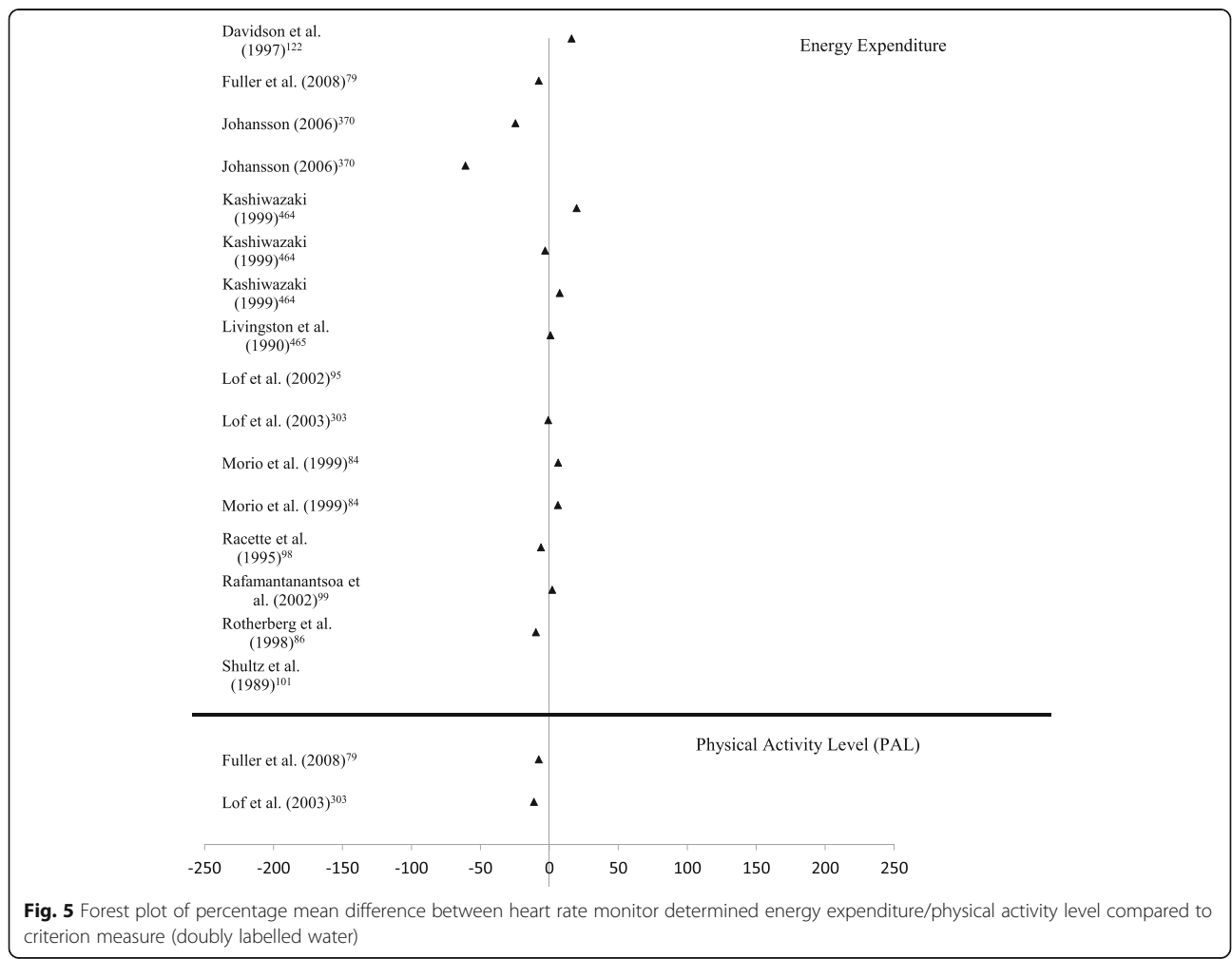
**Heart rate monitors**

**Validity**

*Criterion validity:* All 12 studies that examined the criterion validity of HRMs were unstructured, free-living protocols [80, 85, 87, 96, 99, 100, 102, 123, 304, 371, 465, 466]. The duration of monitoring for HRM ranged from 24 hours [102, 465] to 14 days [96, 371]. Two studies examined the validity of HRM determined physical activity levels (PAL) compared to DLW determined PAL. All remaining articles compared estimates of energy expenditure determined by HRM techniques with DLW determined energy expenditure. The flex heart rate methodology, which distinguishes between activity intensities based on heart rate versus VO<sub>2</sub> calibration curves, were utilised in all studies using individual calibration curves. MPDs between HRM determined energy expenditure and DLW determined energy expenditure ranged from -60.8% - 19.7% across identified studies (Fig. 5). No clear trend for over/under estimation was apparent (MPDs for energy expenditure ranging between

-60.8% - 19.7%). For PAL, a slight trend in underestimation was apparent (-11.1 to -7.6) (Additional file 1: Table S16).

*Concurrent validity:* The concurrent validity of HRM determined energy expenditure [80, 467-470], PAL [80] and PA intensity [146, 174] was examined using a range of measures, including direct/indirect calorimetry [467, 469, 470], activity monitoring [80, 146, 174, 401] and measures of self-reported PA [80, 174, 468] (Additional file 1: Table S17). A slight trend in overestimation of energy expenditure and PAL was observed (Additional file 1: figure S5a). For PA intensities, MPDs were larger and more variable, with MPA underestimated and VPA overestimated. The MPD between HRM determined LIPA and LIPA determined by the Tritrac and MTI activity monitors fell outside the range for the presented forest plot, with values of +306.4% and +367.2%, respectively [146] (Additional file 1: figure S5a). No articles sourced through the data extraction reported on the reliability or responsiveness to change of HRM.



**Fig. 5** Forest plot of percentage mean difference between heart rate monitor determined energy expenditure/physical activity level compared to criterion measure (doubly labelled water)

**Combined sensors**

**Validity**

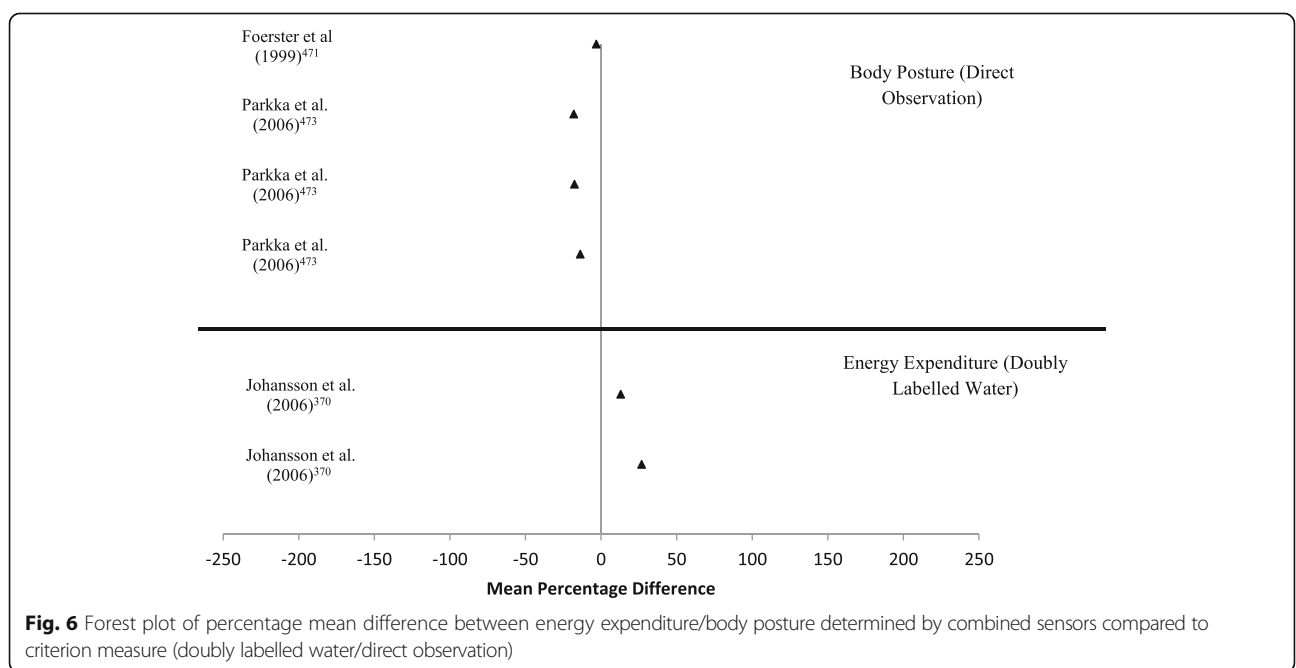
*Criterion validity:* A total of 8 articles were identified that examined the criterion validity of multiple accelerometers [471–474] or accelerometers combined with gyroscopes [475] or HRMs [371, 476, 477]. The included studies had relatively small sample sizes, ranging from 3–31 participants. Studies primarily examined the effectiveness of data synthesis methodologies (i.e. Decision Tree Classification, Artificial Neural Networks, Support Vector Machine learning etc.) to identify specific postures/activities [471–477] or energy expenditure [371, 477]. Time spent in specific body postures/activity types tended to be underestimated from combined sensors when compared to direct observation (-33.3% to -3.2%; Fig. 6). In contrast, energy expenditure was overestimated by combined sensors when compared to DLW in free-living settings (13.0% to 26.8%) (Additional file 1: Table S18) [371].

*Concurrent validity:* Eleven studies examined the validity of combined accelerometry and HRM determined energy expenditure compared to whole room calorimetry [478–480] or indirect calorimetry [400, 477, 481–486] determined energy expenditure. No clear trend for under/overestimation was apparent, with combined sensors appearing to be relatively accurate in estimating energy expenditure when compared to indirect calorimetry in both a structured (-13.8% - 31.1%) and unstructured (0.13%) [485] settings (Additional file 1: Table S19). No articles sourced through the data extraction reported on the reliability or responsiveness to change of combined sensors.

**Discussion**

To the authors’ knowledge, this is the first systematic literature review of reviews to simultaneously examine the methodological effectiveness of the majority of PA measures. The greatest quantity of information was available for self-reported measures of PA (198 data extraction points), followed by activity monitors (179 data extraction points), pedometers (52 data extraction points), HRMs (19 data extraction points) and combined sensors (18 data extraction points).

The criterion validity of measures was determined through the examination of energy expenditure via DLW and by direct observation of steps and PA behaviours. For accelerometry, although variability was lower, a substantial proportion of studies (44/54) underestimated energy expenditure compared to DLW when proprietary algorithms or count-to-activity thresholds were employed. Based on the amended forest plots for the criterion validity of measures of PA, a greater level of variability was apparent for self-reported measures compared to objective measures (Figs. 2–6). Limited data on the criterion validity of HRM and combined sensors determined energy expenditure was available. HRMs tended to underestimate DLW determined energy expenditure, while combined sensors often overestimated energy expenditure. Unfortunately, due to the lack of measures of variability, resulting in the absence of meta-analysis, it was not possible to describe the extent of differences between measures statistically. For step counts, both activity monitors and pedometers achieved high levels of criterion validity. When



**Fig. 6** Forest plot of percentage mean difference between energy expenditure/body posture determined by combined sensors compared to criterion measure (doubly labelled water/direct observation)

comparing the two, pedometers appeared to be less accurate than activity monitors at estimating step count, tending to underestimate steps when compared to direct observation. Activity monitors tended to slightly overestimate distance travelled, while time spent in each activity type (or posture) determined by both activity monitors and combined sensors was slightly underestimated when compared to direct observation (Fig. 3a and Fig. 6). For concurrent validity of all measure of PA, high levels of variability were observed across a wide range of activity behaviours. In particular, high levels of variability were apparent in the estimation of PA intensities, with VPA substantially overestimated in the majority of concurrent validations across all measures. In summary, objective measures are less variable than recall based measures across all behaviours, but high levels of variability across behaviours are still apparent.

For activity monitors and pedometers, acceptable inter-instrument reliability was observed in the majority of studies. Variability for inter-instrument reliability across different activity monitors and pedometers was apparent, with some instruments demonstrating better reliability compared to others. However, a detailed examination of study methodology, device wear locations and activities performed is necessary when interpreting the inter-instrument reliability of pedometers and activity monitors.

A wide range of values were reported for the test-retest reliability of self-reported measures, with apparent trends for reduced levels of test-retest reliability as the duration of recall increased. Researchers must be cognizant of potential differences in test-retest reliability due to duration between administrations and between PA behaviours assessed within each tool when selecting a self-reported measure of PA. Moderate to strong test-retest reliability was observed for activity monitors in free-living environments. However, the reliability of activity monitors attenuated as the duration between measurements increased. As expected, the test-retest reliability of different devices varied, while intensity of activity often had a significant effect. The test-retest reliability of pedometer determined steps in a laboratory setting was high across the majority of speeds, but the reliability appeared to weaken at higher speeds (e.g. 20, 22 and 26 km·h<sup>-1</sup>). Although moderate to strong test-retest reliability of both pedometers and activity monitors were apparent, researchers should be aware of differences between models and devices when selecting a measure for use. Furthermore consideration should be given to the duration between test and retest and the behaviour being assessed when considering test-retest reliability, as although a measure may be reliable for one output, it may not be reliable for all outcomes.

When examining PA in free-living environments, it is essential that sufficient data is gathered to ensure a

reliable estimate is obtained [7, 431]. By determining the inter- and intra-individual variability across days of measurement, researchers can define the number of days of monitoring required to reliably estimate such behaviours. For activity monitors and pedometers, analysis has been conducted to estimate the minimum number of days of measurement required to provide a reliable estimate of PA behaviors. For activity monitors, two days of measurement are recommended for a reliable estimate of steps per day, accelerometer counts per day and intermittent MVPA per day measured, 3 days for a reliable estimate of total PA and time spent in MVPA and 6 days are required for a reliable estimate of continuous 10 minute bouts of MVPA. For pedometers, a minimum of 2-4 days of measurement was required to provide a reliable estimate of steps in older adults, while 2-5 days of measurement was required in adults. These findings highlight the importance of knowing what behaviours are to be examined prior to collecting objective data from free-living environments, to ensure that sufficient information is recorded to provide reliable estimates of the behaviours of interest.

The responsiveness of measures to detect change over time was the least reported property of measures of PA. When evaluating interventions, or indeed evaluating changes in PA behaviours in longitudinal research, it is critical to utilise measures that can detect such changes. Although validity and reliability are requirements for sensitivity/responsiveness to change [5], this does not imply that a measure is responsive to change simply because it is valid and reliable. Responsiveness to change must be evaluated, and not assumed. Currently, the research on the responsiveness to change for all types of PA measurement is at best limited. Substantial investigation into the responsiveness of PA measures to detect change is required to ensure that measures employed in future intervention and longitudinal research can detect meaningful change.

Although the validity, reliability and responsiveness to change are key when selecting a measure of PA and energy expenditure, other factors including feasibility and cost should be considered. For example, wearing several sensors around the body for a short period in a laboratory setting is often quite feasible, but prolonging the wear period for several days may be uncomfortable for participants, while reattachment of sensors may require specific and detailed training. The appropriateness of the measure for use in specific populations is critical. Activity monitors or HRMs may need to be attached to body locations that are visible and may be considered “embarrassing” for certain populations in free-living environments, likely resulting in lower compliance to wear protocols. Finally, while the cost of objective measures have reduced significantly and are now feasible for

inclusion in large scale data collections (i.e. UK Biobank study, HELENA study), worn devices can be expensive to use in large populations, especially if recording needs to be concurrent, requiring 100's or 1000's of devices. Although these issues are often the dominant determinant for researchers when selecting a measure of PA, it is critical that researchers consider selecting the measure with the best validity, reliability and responsiveness to change available to them; a larger dataset with less valid measures may not always be superior to a smaller dataset.

The findings of this review have highlighted the substantial quantity of research which has focused on the validity, reliability and responsiveness to change of measures of PA. A substantial number of review articles have been conducted on the measurement of PA in adult populations. The majority of such reviews were not systematic in nature. Of the systematic reviews articles identified, the methodological quality (as assessed by the AMSTAR quality assessment tool) was relatively poor, with 3 reviews considered low quality, 16 articles considered medium quality and 3 articles considered high quality. An obvious increase in the quantity of research using objective measures of PA over the past number of decades is apparent. Unfortunately, with the enormous quantity of research on the methodological effectiveness of PA measures comes extreme variability in study design, data processing and statistical analysis conducted. Such variability makes comparison between measurement type and specific measurement devices/tools extremely difficult. The sometimes questionable study designs and research questions in some of the existing published literature is a reanalysis of "suitable" data, rather than from a study designed to collect data to answer a specific research questions. The authors propose that to aid researchers in making informed decisions on the best available measure of PA, the development of "best practise" protocols for study design and data collection, analysis and synthesis are required, which can be employed across all measures, providing comparable information that is easy for researchers from outside of the field to digest. The authors also propose that any future undertaking of reviews on the measurement of PA follow best practise, and ensure that the reviews conducted are of the highest possible quality. Such improvements will provide researchers with the best available evidence for making a decision on which measure of PA to employ.

### Strengths and limitations

This review of reviews had limitations that should be taken into account when considering the findings presented here. As this article reviewed existing literature reviews, and due to potential methodological errors within these reviews, it is likely that some relevant literature on the methodological effectiveness for measures of

PA has been overlooked. Additionally, articles that have been published since the publication of each review will also have been overlooked. Due to the quantity of identified articles, and difficulties in contacting primary authors regarding articles published over the last 60 years, the primary data from these articles was not sourced. Although prior research has systematically reviewed the literature for accuracy of measures of PA, and some narrative reviews have compared the methodological effectiveness of different measures of PA, this is the first study to comprehensively examine and collate details on the validity, reliability and responsiveness to change of a range of measures of PA in adult populations. For researchers that are selecting a measure of PA, this will enable the comparison between different measures of PA within one article, rather than having to refer to a wide range of available literature that examines each single measure. Additionally, rather than focusing solely on information presented within each existing review of the literature, the original articles referred to within each review were sought and data was extracted independently.

### Conclusion

In general, objective measures of PA demonstrate less variability in properties of methodological effectiveness than self-report measures. Although no "perfect" tool for the examination of PA exists, it is suggested that researchers aim to incorporate appropriate objective measures, specific to the behaviours of interests, when examining PA in adults in free-living environments. Other criteria beyond methodological effectiveness often influence tool selection, including cost and feasibility. However, researchers must be cognisant of the value of increased methodological effectiveness of any measurement method for PA in adults. Additionally, although a wealth of research exists in relation to the methodological effectiveness of PA measures, it is clear that the development of an appropriate and consistent approach to conducting research and reporting findings in this domain is necessary to enable researchers to easily compare findings across instruments.

### Additional file

**Additional file 1: Table S1.** Criterion validity of self-reported measures of physical activity/energy expenditure. **Table S2.** Concurrent validity of self-reported measures of physical activity/energy expenditure. **Table S3.** Intra-instrument Reliability of self-reported measures of physical activity. **Table S4a.** Test-retest reliability of self-reported measures of physical activity/energy expenditure within a duration of less than or equal to one week. **Table S4b.** Test-retest reliability of self-reported measures of physical activity/energy expenditure within a duration of between 1 week and 4 weeks. **Table S4c.** Test-retest reliability of self-reported measures of physical activity/energy expenditure within a duration of between 4 weeks and 8 weeks. **Table S4d.** Test-retest reliability of self-reported measures of physical activity/energy expenditure within a duration

of between 8 weeks and 1 year. **Table S4e.** Test-retest reliability of self-reported measures of physical activity/energy expenditure within a duration of greater than 1 year. **Table S5.** Sensitivity to change over time of self-reported measures of physical activity/energy expenditure. **Table S6.** Criterion validity of accelerometer activity monitor determined physical activity/energy expenditure. **Table S7.** Concurrent validity of accelerometer/activity monitor determined physical activity/energy expenditure. **Table S8.** Inter-instrument reliability of accelerometer/activity monitor determined physical activity/energy expenditure. **Table S9.** Test-retest reliability of accelerometer/activity monitor determined physical activity/energy expenditure. **Table S10.** Sensitivity to change over time of accelerometer devices. **Table S11.** Details of studies that examined the Criterion Validity of Pedometers. **Table S12.** Details of studies examining the concurrent validity of pedometers. **Table S13.** Details of studies examining inter-instrument reliability in pedometer devices. **Table S14.** Details of studies examining the test-retest reliability of pedometers. **Table S15.** Details of studies examining the sensitivity to change of pedometers. **Table S16.** Details of studies examining the criterion validity of heart rate monitoring devices. **Table S17.** Details of studies examining the concurrent validity of heart rate monitoring devices. **Table S18.** Details of studies examining the criterion validity of combined sensors. **Table S19.** Details of studies examining the concurrent validity of combined sensors. **Figure S1.** PRISMA Checklist. **Figure S2a.** Forest plot of percentage mean difference between self-report measures of energy expenditure compared to energy expenditure from activity logs/diaries. **Figure S2b.** Forest plot of percentage mean difference between self-report measures of energy expenditure and time spent in physical activity compared to other self-report measures of energy expenditure or time spent in physical activity. **Figure S2c.** Forest plot of percentage mean difference between self-report measures of energy expenditure compared to energy expenditure determined from activity monitors. **Figure S2d.** Forest plot of percentage mean difference between self-report measures of time spent in physical activity intensities (Light, Moderate and Moderate-to-Vigorous intensity physical activity) compared to time spent in physical activity intensities determined from activity monitors. **Figure S2e.** Forest plot of percentage mean difference between self-report measures of time spent in physical activity intensities (Vigorous physical activity, Total physical activity, times active, time standing, time stepping) compared to time spent in physical activity intensities determined from activity monitors. **Figure S2f.** Forest plot of percentage mean difference between self-report measures of energy expenditure and time spent in physical activity intensities (Vigorous physical activity, Total physical activity, times active, time standing, time stepping) compared to energy expenditure time spent in physical activity intensities determined from pedometers and heart rate monitors. **Figure S3c.** Forest plot of percentage mean difference between accelerometer/activity monitor determined variables (activity counts, time spent in light intensity physical activity, time spent in moderate intensity physical activity, time spent in moderate-to-vigorous intensity physical activity, time spent in vigorous intensity physical activity, total physical activity, stepping and energy expenditure) compared to an alternative accelerometer/activity monitor. **Figure S3d.** Forest plot of percentage mean difference between accelerometer/activity monitor determined energy expenditure (METs) in light intensity physical activity, moderate intensity physical activity, vigorous intensity physical activity and total physical activity (METs, MJ.d, KJ.h, KJ.kg.min<sup>-1</sup>) compared to estimates from indirect (IC) and whole room calorimetry (WRC). **Figure S3e.** Forest plot of percentage mean difference between accelerometer/activity monitor determined energy expenditure, energy expenditure from light intensity physical activity, moderate intensity physical activity, vigorous intensity physical activity, total physical activity compared to estimates from Heart Rate Monitoring (HRM) and Self-Report (SR) measures. **Figure S3f.** Forest plot of percentage mean difference between accelerometer/activity monitor determined energy expenditure (kcal.min<sup>-1</sup>, kcal.kg.hr<sup>-1</sup>) compared to indirect calorimetry determined energy expenditure (kcal.min<sup>-1</sup>, kcal.kg.hr<sup>-1</sup>). **Figure S3g.** Forest plot of percentage mean difference between accelerometer/activity monitor determined energy expenditure (METs.min<sup>-1</sup>, METs.hr<sup>-1</sup>) compared to indirect calorimetry determined energy expenditure (METs.min<sup>-1</sup>, METs.hr<sup>-1</sup>). **Figure S3h.** Forest plot of percentage mean difference between accelerometer/activity monitor determined total energy expenditure compared to indirect calorimetry determined total energy expenditure. **Figure S3h (cont).** Forest plot of percentage mean difference between accelerometer/activity

monitor determined energy expenditure (kcal.min<sup>-1</sup>, kcal.kg.hr<sup>-1</sup>) compared to indirect calorimetry determined energy expenditure (kcal.min<sup>-1</sup>, kcal.kg.hr<sup>-1</sup>). **Figure S4a.** Forest plot of percentage mean difference between pedometer determined step count/energy expenditure/MVPA compared to concurrent measures (i.e. accelerometry, indirect calorimetry, pedometers). **Figure S5a.** Forest plot of percentage mean difference between heart rate monitor determined energy expenditure/physical activity level/physical activity intensity compared to concurrent measures (accelerometers, self-report, indirect calorimetry) **Figure S6.** Forest plot of percentage mean difference between energy expenditure determined by combined sensors compared to concurrent measure (indirect calorimetry). (DOCX 1304 kb)

#### Acknowledgements

The authors would like to acknowledge the contribution of all members of the DEDIPAC Knowledge Hub consortium for their help and assistance in the completion of this work.

#### Funding

The preparation of this paper was supported by the Determinants of Diet and Physical Activity (DEDIPAC) knowledge hub. This work is supported by the Joint Programming Initiative 'Healthy Diet for a Healthy Life'. This work was funded by the Health Research Board (HRB) of Ireland. Additional funding agencies supporting this work were (in alphabetical order of collaborative Member State): Germany: Project Management Agency in the German Aerospace Centre (PT-DLR); Italy: Ministry of Education, University and Research/ Ministry of Agriculture Food and Forestry Policies; The Netherlands: The Netherlands Organisation for Health Research and Development (ZonMw); Norway: The Research Council of Norway, Division for Society and Health; Poland: The National Centre for Research and Development; The United Kingdom: The Medical Research Council (MRC).

#### Availability of data and materials

All data examined during this study are included in this published article [and its supplementary materials].

#### Author contributions

HvdP, UE and AD were involved in the conception of the review title. KD, RSz, SC, AP, HvdP, UE and AD contributed to the design of the review protocol. KD conducted the database search. Several reviewers independently performed the selection of articles (KD, RSz, JM, RSt, MAM and MT) and examined the titles and abstracts of the identified references to exclude articles out of scope. Any disagreements on study inclusions were resolved through discussions with another reviewer (AD) and a consensus reached. KD, RSz, JM, RSt, MAM and MT assessed the eligible papers, extracted the data, and discussed the findings. KD, RSz, MAM, MM, AP, HvdP, UE and AD drafted the paper and all authors listed reviewed the manuscript and contributed to subsequent drafts. All authors read and approved the final document.

#### Ethics approval and consent to participate

The paper is a literature review on methods, and only employs published research data. There were no human participants, and our University ethics committee does not approve non-human based research.

#### Consent for publication

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

<sup>1</sup>Department of Sport and Health Science, Athlone Institute of Technology, Athlone, Ireland. <sup>2</sup>University School of Physical Education in Poznan, Poznan, Poland. <sup>3</sup>Division of Endocrinology, Diabetology and Metabolism, Department of Internal Medicine, University of Turin, Corso Dogliotti 14,

10126 Torino, Italy. <sup>4</sup>School of Health Science, University of Ulster, Newtownabbey, UK. <sup>5</sup>National Institute for Food and Nutrition Research, Rome, Italy. <sup>6</sup>Department of Public and Occupational Health, VU University Medical Center, EMGO Institute for Health and Care Research, Amsterdam, The Netherlands. <sup>7</sup>Sydney School of Public Health, University of Sydney, Sydney, Australia. <sup>8</sup>Medical Research Council (MRC) Epidemiology Unit, University of Cambridge, Cambridge, UK. <sup>9</sup>The Department of Sport Medicine, Norwegian School of Sport Sciences, Oslo, Norway. <sup>10</sup>Department of Physical Education and Sport Sciences, Health Research Institute, University of Limerick, Limerick, Ireland.

Received: 28 June 2017 Accepted: 18 December 2017

Published online: 08 February 2018

## References

- Kohl HW, Craig CL, Lambert EV, Inoue S, Alkandari JR, Leetongin G, Kahlmeier S. Group LPASW: The pandemic of physical inactivity: global action for public health. *Lancet*. 2012;380:294–305.
- Haskell WL, Lee I-M, Pate RR, Powell KE, Blair SN, Franklin BA, Macera CA, Heath GW, Thompson PD, Bauman A. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Circulation*. 2007;116:1081–93.
- Owen N, Healy GN, Matthews CE, Dunstan DW. Too much sitting: the population-health science of sedentary behavior. *Exerc Sport Sci Rev*. 2010;38:105–13.
- Shephard RJ, Aoyagi Y. Measurement of human energy expenditure, with particular reference to field studies: an historical perspective. *Eur J Appl Physiol*. 2012;112:2785–815.
- Warren JM, Ekelund U, Besson H, Mezzani A, Geladas N, Vanhees L. Assessment of physical activity - a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. *Eur J Prev Cardiol*. 2010;17:127–39.
- Dishman RK, Washburn RA, Schoeller DA. Measurement of physical activity. *Quest*. 2001;53:295–309.
- Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc*. 2005;37:S531–43.
- Adams SA, Matthews CE, Ebbeling CB, Moore CG, Cunningham JE, Fulton J, Hebert JR. The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol*. 2005;161:389–98.
- Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. *Res Q Exerc Sport*. 2000;71:1–14.
- Neilson HK, Robson PJ, Friedenreich CM, Cizmadzi I. Estimating activity energy expenditure: how valid are physical activity questionnaires? *Am J Clin Nutr*. 2008;87:279–91.
- van Poppel MNM, Chinapaw MJM, Mokkink LB, van Mechelen W, Terwee CB. Physical Activity Questionnaires for Adults A Systematic Review of Measurement Properties. *Sports Med*. 2010;40:565–600.
- Ainslie PN, Reilly T, Westerterp KR. Estimating Human Energy Expenditure: A Review of Techniques with Particular Reference to Doubly Labelled Water. *Sports Med*. 2003;33:683–98.
- Bort-Roig J, Gilson N, Puig-Ribera A, Contreras R, Trost S. Measuring and Influencing Physical Activity with Smartphone Technology: A Systematic Review. *Sports Med*. 2014;44:671–86.
- Freedson P, Miller K. Objective monitoring of physical activity using motion sensors and heart rate. *Res Q Exerc Sport*. 2000;71:21–9.
- Tudor-Locke CE, Myers AM. Methodological considerations for researchers and practitioners using pedometers to measure physical (ambulatory) activity. *Res Q Exerc Sport*. 2001;72:1–12.
- Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber SC, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *Int J Behav Nutr Phys Act*. 2008;5:56.
- Bassett DR Jr, Mahar MT, Rowe DA, Morrow JR Jr. Walking and measurement. *Med Sci Sports Exerc*. 2008;40:S529–36.
- Kim Y, Park I, Kang M. Convergent validity of the international physical activity questionnaire (IPAQ): meta-analysis. *Public Health Nutr*. 2013;16:440–52.
- Kwak L, Proper KI, Hagströmer M, Sjöström M. The repeatability and validity of questionnaires assessing occupational physical activity - a systematic review. *Scand J Work Environ Health*. 2011;37:6–29.
- Bassett DR Jr. Validity and reliability issues in objective monitoring of physical activity. *Res Q Exerc Sport*. 2000;71:S30–6.
- Bonomi AG, Westerterp KR. Advances in physical activity monitoring and lifestyle interventions in obesity: a review. *Int J Obes*. 2012;36:167–77.
- Washburn RA. Assessment of physical activity in older adults. *Res Q Exerc Sport*. 2000;71:79–88.
- Pennathur A, Magham R, Contreras LR, Dowling W. Daily living activities in older adults: Part I - a review of physical activity and dietary intake assessment methods. *Int J Ind Ergon*. 2003;32:389–404.
- Meyer K, Stolz C, Rott C, Laederach-Hofmann K. Physical activity assessment and health outcomes in old age: how valid are dose-response relationships in epidemiologic studies? *Eur Rev Aging Phys Act*. 2009;6:7–17.
- Murphy SL. Review of physical activity measurement using accelerometers in older adults: considerations for research design and conduct. *Prev Med*. 2009;48:108–14.
- Kowalski K, Rhodes R, Naylor PJ, Tuokko H, MacDonald S. Direct and indirect measurement of physical activity in older adults: a systematic review of the literature. *Int J Behav Nutr Phys Act*. 2012;9:148.
- Garatachea N, Torres Luque G, Gonzalez Gallego J. Physical activity and energy expenditure measurements using accelerometers in older adults. *Nutr Hosp*. 2010;25:224–30.
- Gorman E, Hanson HM, Yang PH, Khan KM, Liu-Ambrose T, Ashe MC. Accelerometry analysis of physical activity and sedentary behavior in older adults: a systematic review and data analysis. *Eur Rev Aging Phys Act*. 2014;11:35–49.
- Forsen L, Loland NW, Vuillemin A, Chinapaw MJM, van Poppel MNM, Mokkink LB, van Mechelen W, Terwee CB. Self-Administered Physical Activity Questionnaires for the Elderly A Systematic Review of Measurement Properties. *Sports Med*. 2010;40:601–23.
- Cheung VH, Gray L, Karunanithi M. Review of Accelerometry for Determining Daily Activity Among Elderly Patients. *Arch Phys Med Rehabil*. 2011;92:998–1014.
- Helmerhorst HJ, 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;9:103.
- Kowalski K, Rhodes R, Naylor PJ, Tuokko H, MacDonald S. Direct and indirect measurement of physical activity in older adults: a systematic review of the literature. *Int J Behav Nutr Phys Act*. 2012;9:148.
- Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. *Obesity*. 2007;15:2371–9.
- Reiser LM, Schlenk EA. Clinical use of physical activity measures. *J Am Acad Nurse Pract*. 2009;21:87–94.
- Centre for reviews dissemination. Systematic reviews: CRD's guidance for undertaking reviews in health care: Centre for Reviews and Dissemination; 2009.
- Melanson EL Jr, Freedson PS, Blair S. Physical activity assessment: a review of methods. *Crit Rev Food Sci Nutr*. 1996;36:385–96.
- Duncan MJ, Badland HM, Mummery WK. Applying GPS to enhance understanding of transport-related physical activity. *J Sci Med Sport*. 2009;12:549–56.
- Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007;7:10.
- Berlin JE, Storti KL, Brach JS. Using activity monitors to measure physical activity in free-living conditions. *Phys Ther*. 2006;86:1137–45.
- Butte NF, Ekelund U, Westerterp KR. Assessing physical activity using wearable monitors: measures of physical activity. *Med Sci Sports Exerc*. 2012;44:5–12.
- Chen KY, Bassett DR. The technology of accelerometry-based activity monitors: Current and future. *Med Sci Sports Exerc*. 2005;37:490–S500.
- Corder K, Brage S, Ekelund U. Accelerometers and pedometers: methodology and clinical application. *Curr Opin Clin Nutr Metab Care*. 2007;10:597–603.
- Davidson M, de Morton N. A systematic review of the Human Activity Profile. *Clin Rehabil*. 2007;21:151–62.
- DeLany JP. Measurement of energy expenditure. *Pediatric Blood and Cancer*. 2012;58:129–34.
- Haskell WL, Kiernan M. Methodologic issues in measuring physical activity and physical fitness when evaluating the role of dietary supplements for physically active people. *Am J Clin Nutr*. 2000;72:541s–50s.

46. Lamonte MJ, Ainsworth BE. Quantifying energy expenditure and physical activity in the context of dose response. *Med Sci Sports Exerc.* 2001;33:5370–8.
47. 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.
48. Levine JA. Measurement of energy expenditure. *Public Health Nutr.* 2005;8:1123–32.
49. Liu SP, Gao RX, Freedson PS. Computational Methods for Estimating Energy Expenditure in Human Physical Activities. *Med Sci Sports Exerc.* 2012;44:2138–46.
50. Lowe SA, Ólaighin G. Monitoring human health behaviour in one's living environment: A technological review. *Med Eng Phys.* 2014;36:147–68.
51. Mathie MJ, Coster AC, Lovell NH, Celler BG. Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. *Physiol Meas.* 2004;25:R1–20.
52. Matthews CE. Calibration of accelerometer output for adults. *Med Sci Sports Exerc.* 2005;37:512–522.
53. Pedišić Ž, Bauman A. Accelerometer-based measures in physical activity surveillance: current practices and issues. *Br J Sports Med.* 2014;49:219–23.
54. Pierannunzi C, Hu SS, Balluz L. A systematic review of publications assessing reliability and validity of the Behavioral Risk Factor Surveillance System (BRFSS), 2004–2011. *BMC Med Res Methodol.* 2013;13:49.
55. Plasqui G, Bonomi AG, Westerterp KR. Daily physical activity assessment with accelerometers: new insights and validation studies. *Obes Rev.* 2013;14:451–62.
56. Reilly JJ, Penpraze V, Hislop J, Davies G, Grant S, Paton JY. Objective measurement of physical activity and sedentary behaviour: review with new data. *Arch Dis Child.* 2008;93:614–9.
57. Ridgers ND, Fairclough S. Assessing free-living physical activity using accelerometry: Practical issues for researchers and practitioners. *Eur J Sport Sci.* 2011;11:205–13.
58. Schutz Y, Weinsier RL, Hunter GR. Assessment of free-living physical activity in humans: an overview of currently available and proposed new measures. *Obes Res.* 2001;9:368–79.
59. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med.* 2003;37:197–206.
60. Strath SJ, Kaminsky LA, Ainsworth BE, Ekelund U, Freedson PS, Gary RA, Richardson CR, Smith DT, Swartz AM, Council Clinical C, Council Cardiovasc Stroke N. Guide to the Assessment of Physical Activity: Clinical and Research Applications A Scientific Statement From the American Heart Association. *Circulation.* 2013;128:2259–79.
61. Tudor-Locke C, Rowe DA. Using cadence to study free-living ambulatory behaviour. *Sports Med.* 2012;42:381–98.
62. Tudor-Locke C, Williams JE, Reis JP, Pluto D. Utility of pedometers for assessing physical activity: convergent validity. *Sports Med.* 2002;32:795–808.
63. Tudor-Locke C, Williams JE, Reis JP, Pluto D. Utility of pedometers for assessing physical activity: construct validity. *Sports Med.* 2004;34:281–91.
64. Valanou EM, Bamia C, Trichopoulou A. Methodology of physical-activity and energy-expenditure assessment: A review. *J Public Health.* 2006;14:58–65.
65. Van Remoortel H, Giavedoni S, Raste Y, Burtin C, Louvaris Z, Gimeno-Santos E, Langer D, Glendenning A, Hopkinson NS, Vogiatzis I, et al. Validity of activity monitors in health and chronic disease: a systematic review. *Int J Behav Nutr Phys Act.* 2012;9:84.
66. Vanhees L, Lefevre J, Philippaerts R, Martens M, Huygens W, Troosters T, Beunen G. How to assess physical activity? How to assess physical fitness? *Eur J Prev Cardiol.* 2005;12:102–14.
67. Washburn RA, Heath GW, Jackson AW. Reliability and validity issues concerning large-scale surveillance of physical activity. *Res Q Exerc Sport.* 2000;71:104–13.
68. Welk GJ. Principles of design and analyses for the calibration of accelerometry-based activity monitors. *Med Sci Sports Exerc.* 2005;37:501–11.
69. Westerterp KR. Assessment of physical activity: a critical appraisal. *Eur J Appl Physiol.* 2009;105:823–8.
70. Westerterp KR, Plasqui G. Physical activity and human energy expenditure. *Curr Opin Clin Nutr Metab Care.* 2004;7:607–13.
71. Yang CC, Hsu YL. A review of accelerometry-based wearable motion detectors for physical activity monitoring. *Sensors.* 2010;10:7772–88.
72. Andrew NE, Gabbe BJ, Wolfe R, Cameron PA. Evaluation of Instruments for Measuring the Burden of Sport and Active Recreation Injury. *Sports Med.* 2010;40:141–61.
73. Adams SA, Matthews CE, Ebbeling CB, Moore CG, Cunningham JE, Fulton J, et al. The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol.* 2005;161:389–98.
74. Barnard J, Tapsell LC, Davies P, Brenninger V, Storlien L. Relationship of high energy expenditure and variation in dietary intake with reporting accuracy on 7 day food records and diet histories in a group of healthy adult volunteers. *Eur J Clin Nutr.* 2002;56:358–67.
75. Besson H, Brage S, Jakes RW, Ekelund U, Wareham NJ. Estimating physical activity energy expenditure, sedentary time, and physical activity intensity by self-report in adults. *Am J Clin Nutr.* 2010;91:106–14.
76. Bonnefoy M, Normand S, Pachiaudi C, Lacour JR, Laville M, Kostka T. Simultaneous validation of ten physical activity questionnaires in older men: a doubly labeled water study. *J Am Geriatr Soc.* 2001;49:28–35.
77. Colbert LH, Matthews CE, Havighurst TC, Kim K, Schoeller DA. Comparative validity of physical activity measures in older adults. *Med Sci Sports Exerc.* 2011;43:867–76.
78. Conway JM, Irwin ML, Ainsworth BE. Estimating energy expenditure from the Minnesota Leisure Time Physical Activity and Tecumseh Occupational Activity questionnaires—a doubly labeled water validation. *J Clin Epidemiol.* 2002;55:392–9.
79. Conway JM, Seale JL, Jacobs DR, Irwin ML, Ainsworth BE. Comparison of energy expenditure estimates from doubly labeled water, a physical activity questionnaire, and physical activity records. *Am J Clin Nutr.* 2002;75:519–25.
80. Fuller Z, Horgan G, O'Reilly L, Ritz P, Milne E, Stubbs R. Comparing different measures of energy expenditure in human subjects resident in a metabolic facility. *Eur J Clin Nutr.* 2008;62:560–9.
81. Goran MI, Poehlman ET. Total energy expenditure and energy requirements in healthy elderly persons. *Metabolism.* 1992;41:744–53.
82. Irwin ML, Ainsworth BE, Conway JM. Estimation of energy expenditure from physical activity measures: determinants of accuracy. *Obes Res.* 2001;9:517–25.
83. Maddison R, Mhurchu CN, Jiang Y, Vander Hoorn S, Rodgers A, Lawes CM, Rush E. International physical activity questionnaire (IPAQ) and New Zealand physical activity questionnaire (NZPAQ): A doubly labelled water validation. *Int J Behav Nutr Phys Act.* 2007;4:62.
84. Mase LC, Fulton JE, Watson KL, Mahar MT, Meyers MC, Wong WW. Influence of body composition on physical activity validation studies using doubly labeled water. *J Appl Physiol.* 2004;96:1357–64.
85. Morio B, Ritz P, Verdier E, Montaurier C, Beaufriere B, Vermorel M. Critical evaluation of the factorial and heart-rate recording methods for the determination of energy expenditure of free-living elderly people. *Br J Nutr.* 1997;78:709–22.
86. Paul DR, Rhodes DG, Kramer M, Baer DJ, Rumpel WW. Validation of a food frequency questionnaire by direct measurement of habitual ad libitum food intake. *Am J Epidemiol.* 2005;162:806–14.
87. Rothenberg E, Bosaeus I, Lernfelt B, Landahl S, Steen B. Energy intake and expenditure: validation of a diet history by heart rate monitoring, activity diary and doubly labeled water. *Eur J Clin Nutr.* 1998;52:832–8.
88. Rush E, Valencia ME, Plank LD. Validation of a 7-day physical activity diary against doubly-labelled water. *Ann Hum Biol.* 2008;35:416–21.
89. Schulz LO, Harper IT, Smith CJ, Kriska AM, Ravussin E. Energy intake and physical activity in Pima Indians: comparison with energy expenditure measured by doubly-labeled water. *Obes Res.* 1994;2:541–8.
90. Starling RD, Toth MJ, Matthews DE, Poehlman ET. Energy Requirements and Physical Activity of Older Free-Living African-Americans: A Doubly Labeled Water Study 1. *J Clin Endocrinol Metab.* 1998;83:1529–34.
91. Staten LK, Taren DL, Howell WH, Tobar M, Poehlman ET, Hill A, Reid PM, Ritenbaugh C. Validation of the Arizona Activity Frequency Questionnaire using doubly labeled water. *Med Sci Sports Exerc.* 2001;33:1959–67.
92. Koebernick C, Wagner K, Thielecke F, Moeseneder J, Hoehne A, Franke A, Meyer H, Garcia A, Trippo U, Zunft H. Validation of a simplified physical activity record by doubly labeled water technique. *Int J Obes.* 2005;29:302–9.
93. Leenders N, Sherman WM, Nagaraja H, Kien CL. Evaluation of methods to assess physical activity in free-living conditions. *Med Sci Sports Exerc.* 2001;33:1233–40.
94. Livingstone M, Strain J, Prentice A, Coward W, Nevin G, Barker M, Hickey R, McKenna P, Whitehead R. Potential contribution of leisure activity to the energy expenditure patterns of sedentary populations. *Br J Nutr.* 1991;65:145–55.
95. Lof M, Forsum E. Validation of energy intake by dietary recall against different methods to assess energy expenditure. *J Hum Nutr Diet.* 2004;17:471–80.



96. Lof M, Hannestad U, Forsum E. Assessing physical activity of women of childbearing age. Ongoing work to develop and evaluate simple methods. *Food Nutr Bull.* 2002;23:30–3.
97. Mahabir S, Baer DJ, Giffen C, Clevidence BA, Campbell WS, Taylor PR, Hartman TJ. Comparison of energy expenditure estimates from 4 physical activity questionnaires with doubly labeled water estimates in postmenopausal women. *Am J Clin Nutr.* 2006;84:230–6.
98. Philippaerts R, Westerterp K, Lefevre J. Doubly labelled water validation of three physical activity questionnaires. *Int J Sports Med.* 1999;20:284–9.
99. Racette SB, Schoeller DA, Kushner RF. Comparison of heart rate and physical activity recall with doubly labeled water in obese women. *Med Sci Sports Exerc.* 1995;27:126–33.
100. Rafamantanantsoa HH, Ebine N, Yoshioka M, Higuchi H, Yoshitake Y, Tanaka H, Saitoh S, Jones PJH. Validation of three alternative methods to measure total energy expenditure against the doubly labeled water method for older Japanese men. *J Nutr Sci Vitaminol (Tokyo).* 2002;48:517–23.
101. Reilly J, Lord A, Bunker V, Prentice A, Coward W, Thomas A, Briggs R. Energy balance in healthy elderly women. *Br J Nutr.* 1993;69:21–7.
102. Schulz S, Westerterp KR, Brück K. Comparison of energy expenditure by the doubly labeled water technique with energy intake, heart rate, and activity recording in man. *Am J Clin Nutr.* 1989;49:1146–54.
103. Schuit AJ, Schouten EG, Westerterp KR, Saris WH. Validity of the Physical Activity Scale for the Elderly (PASE): according to energy expenditure assessed by the doubly labeled water method. *J Clin Epidemiol.* 1997;50:541–6.
104. Seale JL, Klein G, Friedmann J, Jensen GL, Mitchell DC, Smiciklas-Wright H. Energy expenditure measured by doubly labeled water, activity recall, and diet records in the rural elderly. *Nutrition.* 2002;18:568–73.
105. Starling RD, Matthews DE, Ades PA, Poehlman ET. Assessment of physical activity in older individuals: a doubly labeled water study. *J Appl Physiol.* 1999;86:2090–6.
106. Walsh MC, Hunter GR, Sirikul B, Gower BA. Comparison of self-reported with objectively assessed energy expenditure in black and white women before and after weight loss. *Am J Clin Nutr.* 2004;79:1013–9.
107. Washburn RA, Jacobsen DJ, Sonko BJ, Hill JO, Donnelly JE. The validity of the Stanford Seven-Day Physical Activity Recall in young adults. *Med Sci Sports Exerc.* 2003;35:1374–80.
108. Aadahl M, Jørgensen T. Validation of a new self-report instrument for measuring physical activity. *Med Sci Sports Exerc.* 2003;35:1196–202.
109. Ainsworth B, Bassett DR Jr, Strath SJ, Swartz AM, O'Brien WL, Thompson RW, Jones DA, Macera CA, Kimsey CD. Comparison of three methods for measuring the time spent in physical activity. *Med Sci Sports Exerc.* 2000;32: S457–64.
110. Ainsworth BE, Leon AS, Richardson MT, Jacobs DR, Paffenbarger R. Accuracy of the college alumnus physical activity questionnaire. *J Clin Epidemiol.* 1993;46:1403–11.
111. Ainsworth BE, Richardson MT, Jacobs DR, Leon AS, Sternfeld B. Accuracy of recall of occupational physical activity by questionnaire. *J Clin Epidemiol.* 1999;52:219–27.
112. Arroll B, Jackson R, Beaglehole R. Validation of a three-month physical activity recall questionnaire with a seven-day food intake and physical activity diary. *Epidemiology.* 1991;2:296–9.
113. Baranowski T, Dworkin RJ, Cieslik CJ, Hooks P, Clearman DR, Ray L, Dunn JK, Nader PR. Reliability and validity of self report of aerobic activity: Family Health Project. *Res Q Exerc Sport.* 1984;55:309–17.
114. Bassett DR Jr, Cureton AL, Ainsworth BE. Measurement of daily walking distance-questionnaire versus pedometer. *Med Sci Sports Exerc.* 2000;32: 1018–23.
115. Boon RM, Hamlin MJ, Steel GD, Ross JJ. Validation of the New Zealand physical activity questionnaire (NZPAQ-LF) and the international physical activity questionnaire (IPAQ-LF) with accelerometry. *Br J Sports Med.* 2008; 44:741–6.
116. Brown WJ, Burton NW, Marshall AL, Miller YD. Reliability and validity of a modified self-administered version of the Active Australia physical activity survey in a sample of mid-age women. *Aust N Z J Public Health.* 2008;32: 535–41.
117. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. *J Phys Act Health.* 2009;6: 790–804.
118. Bulley C, Donaghy M, Payne A, Mutrie N. Validation and modification of the Scottish Physical Activity Questionnaire for use in a female student population. *Int J Health Promot Educ.* 2005;43:117–24.
119. Busse M, Van Deursen R, Wiles CM. Real-life step and activity measurement: reliability and validity. *J Med Eng Technol.* 2009;33:33–41.
120. Carter-Nolan PL, Adams-Campbell LL, Makambi K, Lewis S, Palmer JR, Rosenberg L. Validation of physical activity instruments: Black Women's Health Study. *Ethn Dis.* 2006;16:943–7.
121. Cartmel B, Moon T. Comparison of two physical activity questionnaires, with a diary, for assessing physical activity in an elderly population. *J Clin Epidemiol.* 1992;45:877–83.
122. Craig CL, Marshall AL, Sjostrom M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, Sallis JF, Oja P. International physical activity questionnaire: 12-country reliability and validity. *Med Sci Sports Exerc.* 2003;35:1381–95.
123. Davidson L, McNeill G, Haggarty P, Smith JS, Franklin MF. Free-living energy expenditure of adult men assessed by continuous heart-rate monitoring and doubly-labelled water. *Br J Nutr.* 1997;78:695–708.
124. De Cocker KA, De Bourdeaudhuij IM, Cardon GM. What do pedometer counts represent? A comparison between pedometer data and data from four different questionnaires. *Public Health Nutr.* 2009;12:74–81.
125. Deng H, Macfarlane D, Thomas G, Lao X, Jiang C, Cheng K, Lam T. Reliability and validity of the IPAQ-Chinese: the Guangzhou Biobank Cohort study. *Med Sci Sports Exerc.* 2008;40:303–7.
126. Dinger MK. Reliability and convergent validity of the National College Health Risk Behavior Survey physical activity items. *J Health Educ.* 2003;34: 162–6.
127. Dinger MK, Behrens TK, Han JL. Validity and reliability of the International Physical Activity Questionnaire in college students. *J Health Educ.* 2006;37: 337–43.
128. Dishman R, Steinhardt M. Reliability and concurrent validity for a 7-d re-call of physical activity in college students. *Med Sci Sports Exerc.* 1988;20:14–25.
129. Duncan GE, Sydeman SJ, Perri MG, Limacher MC, Martin AD. Can sedentary adults accurately recall the intensity of their physical activity? *Prev Med.* 2001;33:18–26.
130. Ekelund U, Sepp H, Brage S, Becker W, Jakes R, Hennings M, Wareham NJ. Criterion-related validity of the last 7-day, short form of the International Physical Activity Questionnaire in Swedish adults. *Public Health Nutr.* 2006;9: 258–65.
131. Fjeldsoe B, Marshall A, Miller Y. Measurement properties of the Australian women's activity survey. *Med Sci Sports Exerc.* 2009;41:1020–33.
132. Friedenreich CM, Courneya KS, Neilson HK, Matthews CE, Willis G, Irwin M, Troiano R, Ballard-Barbash R. Reliability and validity of the past year total physical activity questionnaire. *Am J Epidemiol.* 2006;163:959–70.
133. Grimm EK, Swartz AM, Hart T, Miller NE, Strath SJ. Comparison of the IPAQ-Short Form and accelerometry predictions of physical activity in older adults. *J Aging Phys Act.* 2012;20:64–79.
134. Hagstromer M, Ainsworth BE, Oja P, Sjostrom M. Comparison of a subjective and an objective measure of physical activity in a population sample. *J Phys Act Health.* 2010;7:541–50.
135. Hallal PC, Simoes E, Reichert FF, Azevedo MR, Ramos LR, Pratt M, Brownson RC. Validity and reliability of the telephone-administered International Physical Activity Questionnaire in Brazil. *J Phys Act Health.* 2010;7:402–9.
136. Hallal PC, Victora CG, Wells JCK, Rdc L, Valle NJ. Comparison of short and full-length International Physical Activity Questionnaires. *J Phys Act Health.* 2004;1:227–34.
137. Hekler EB, Buman MP, Haskell WL, Conway TL, Cain KL, Sallis JF, Saelens BE, Frank LD, Kerr J, King AC. Reliability and validity of CHAMPS self-reported sedentary-to-vigorous intensity physical activity in older adults. *J Phys Act Health.* 2012;9:225–36.
138. Hopkins WG, Wilson NC, Russell DG. Validation of the physical activity instrument for the Life in New Zealand national survey. *Am J Epidemiol.* 1991;133:73–82.
139. Hurtig-Wennlöf A, Hagströmer M, Olsson LA. The International Physical Activity Questionnaire modified for the elderly: aspects of validity and feasibility. *Public Health Nutr.* 2010;13:1847–54.
140. Kurtze N, Rangul V, Hustvedt B-E. Reliability and validity of the international physical activity questionnaire in the Nord-Trøndelag health study (HUNT) population of men. *BMC Med Res Methodol.* 2008;8:63.
141. Lee PH, Yu Y, McDowell I, Leung GM, Lam T, Stewart SM. Performance of the international physical activity questionnaire (short form) in subgroups of the Hong Kong chinese population. *Int J Behav Nutr Phys Act.* 2011;8:1479–5868.
142. Leenders NY, Sherman WM, Nagaraja H. Comparisons of four methods of estimating physical activity in adult women. *Med Sci Sports Exerc.* 2000;32: 1320–6.

143. Lemmer JT, Ivey FM, Ryan AS, Martel GF, Hurlbut DE, Metter JE, Fozard JL, Fleg JL, Hurley BF. Effect of strength training on resting metabolic rate and physical activity: age and gender comparisons. *Med Sci Sports Exerc.* 2001;33:532–41.
144. Macfarlane D, Chan A, Cerin E. Examining the validity and reliability of the Chinese version of the International Physical Activity Questionnaire, long form (IPAQ-LC). *Public Health Nutr.* 2011;14:443–50.
145. Macfarlane DJ, Lee CC, Ho EY, Chan K, Chan DT. Reliability and validity of the Chinese version of IPAQ (short, last 7 days). *J Sci Med Sport.* 2007;10:45–51.
146. Macfarlane DJ, Lee CC, Ho EY, Chan K-L, Chan D. Convergent validity of six methods to assess physical activity in daily life. *J Appl Physiol.* 2006;101:1328–34.
147. Mäder U, Martin BW, Schutz Y, Marti B. Validity of four short physical activity questionnaires in middle-aged persons. *Med Sci Sports Exerc.* 2006;38:1255–66.
148. Masse LC, Eason KE, Tortolero SR, Kelder SH. Comparing participants' rating and compendium coding to estimate physical activity intensities. *Meas Phys Educ Exerc Sci.* 2005;9:1–20.
149. Matthews CE, Ainsworth BE, Hanby C, Pate RR, Addy C, Freedson PS, Jones DA, Macera CA. Development and testing of a short physical activity recall questionnaire. *Med Sci Sports Exerc.* 2005;37:986–94.
150. Matthews CE, Freedson PS. Field trial of a three-dimensional activity monitor: comparison with self report. *Med Sci Sports Exerc.* 1995;27:1071–8.
151. Matthews CE, Keadle SK, Sampson J, Lyden K, Bowles HR, Moore SC, Libertine A, Freedson PS, Fowke JH. Validation of a previous-day recall measure of active and sedentary behaviors. *Med Sci Sports Exerc.* 2013;45:1629–38.
152. Matthews CE, Shu X-O, Yang G, Jin F, Ainsworth BE, Liu D, Gao Y-T, Zheng W. Reproducibility and validity of the Shanghai Women's Health Study physical activity questionnaire. *Am J Epidemiol.* 2003;158:1114–22.
153. Meriwether RA, McMahon PM, Islam N, Steinmann WC. Physical activity assessment: validation of a clinical assessment tool. *Am J Prev Med.* 2006;31:484–91.
154. Moy KL, Scragg RK, McLean G, Carr H. The New Zealand Physical Activity Questionnaires: validation by heart-rate monitoring in a multiethnic population. *J Phys Act Health.* 2008;5:545–61.
155. EEK N, SAG N, Wu Y, Salim A, Tai ES, Lee J, Van Dam RM. Validity of the International Physical Activity Questionnaire and the Singapore Prospective Study Program physical activity questionnaire in a multiethnic urban Asian population. *BMC Med Res Methodol.* 2011;11:141.
156. Nicaise V, Marshall S, Ainsworth BE. Domain-specific physical activity and self-report bias among low-income Latinas living in San Diego County. *J Phys Act Health.* 2011;8:881–90.
157. Norman A, Belloc R, Bergström A, Wolk A. Validity and reproducibility of self-reported total physical activity—differences by relative weight. *Int J Obes Relat Metab Disord.* 2001;25:682–8.
158. Resnick B, King A, Riebe D, Ory M. Measuring Physical Activity in Older Adults Use of the Community Health Activities Model Program for Seniors Physical Activity Questionnaire and the Yale Physical Activity Survey in Three Behavior Change Consortium Studies. *West J Nurs Res.* 2008;30:673–89.
159. Richardson MT, Ainsworth BE, Jacobs DR, Leon AS. Validation of the Stanford 7-day recall to assess habitual physical activity. *Ann Epidemiol.* 2001;11:145–53.
160. Richardson MT, Leon AS, Jacobs DR Jr, Ainsworth BE, Serfass R. Ability of the Caltrac accelerometer to assess daily physical activity levels. *J Cardiopulm Rehabil Prev.* 1995;15:107–13.
161. Roman-Viñas B, Serra-Majem L, Hagströmer M, Ribas-Barba L, Sjöström M, Segura-Cardona R. International physical activity questionnaire: reliability and validity in a Spanish population. *Eur J Sport Sci.* 2010;10:297–304.
162. Saglam M, Arikani H, Savci S, Inal-Ince D, Bosnak-Guclu M, Karabulut E, Tokgozoglu L. International physical activity questionnaire: reliability and validity of the Turkish version 1. *Percept Mot Skills.* 2010;111:278–84.
163. Schmidt MD, Freedson PS, Chasan-Taber L. Estimating physical activity using the CSA accelerometer and a physical activity log. *Med Sci Sports Exerc.* 2003;35:1605–11.
164. Sidney S, Jacobs DR, Haskell WL, Armstrong MA, Dimicco A, Oberman A, Savage PJ, Slattery ML, Sternfeld B, Van Hom L. Comparison of two methods of assessing physical activity in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. *Am J Epidemiol.* 1991;133:1231–45.
165. Singh PN, Fraser GE, Knutsen SF, Lindsted KD, Bennett HW. Validity of a physical activity questionnaire among African-American Seventh-day Adventists. *Med Sci Sports Exerc.* 2001;33:468–75.
166. Sirard J, Melanson EL, Li L, Freedson PS. Field evaluation of the Computer Science and Applications, Inc. physical activity monitor. *Med Sci Sports Exerc.* 1998;32:695–700.
167. Smith BJ, Marshall AL, Huang N. Screening for physical activity in family practice: evaluation of two brief assessment tools. *Am J Prev Med.* 2005;29:256–64.
168. Sobngwi E, Mbanya JCN, Unwin NC, Aspray TJ, Alberti K. Development and validation of a questionnaire for the assessment of physical activity in epidemiological studies in Sub-Saharan Africa. *Int J Epidemiol.* 2001;30:1361–8.
169. Stel VS, Smit JH, Pluijm SM, Visser M, Deeg DJ, Lips P. Comparison of the LASA Physical Activity Questionnaire with a 7-day diary and pedometer. *J Clin Epidemiol.* 2004;57:252–8.
170. Strath SJ, Bassett DR, Swartz AM. Comparison of the college alumnus questionnaire physical activity index with objective monitoring. *Ann Epidemiol.* 2004;14:409–15.
171. Strath SJ, Bassett DR Jr, Ham SA, Swartz AM. Assessment of physical activity by telephone interview versus objective monitoring. *Med Sci Sports Exerc.* 2003;35:2112–8.
172. Suzuki I, Kawakami N, Shimizu H. Reliability and validity of a questionnaire for assessment of energy expenditure and physical activity in epidemiological studies. *J Epidemiol.* 1998;8:152–9.
173. Tehard B, Saris WH, Astrup A, Martinez JA, Taylor MA, Barbe P, Richterova B, Guy-Grand B, Sorensen T, Oppert J. Comparison of two physical activity questionnaires in obese subjects: the NUGENOB study. *Med Sci Sports Exerc.* 2005;37:1535–41.
174. Theou O, Jakobi JM, Vandervoort AA, Jones GR. A comparison of physical activity (PA) assessment tools across levels of frailty. *Arch Gerontol Geriatr.* 2012;54:e307–14.
175. Thuy AB, Blizzard L, Schmidt M, Hung Luc P, Magnussen C, Dwyer T. Reliability and validity of the global physical activity questionnaire in Vietnam. *J Phys Act Health.* 2010;7:410–8.
176. Timperio A, Salmon J, Crawford D. Validity and reliability of a physical activity recall instrument among overweight and non-overweight men and women. *J Sci Med Sport.* 2003;6:477–91.
177. Timperio A, Salmon J, Rosenberg M, Bull FC. Do logbooks influence recall of physical activity in validation studies? *Med Sci Sports Exerc.* 2004;36:1181–6.
178. Tomioka K, Iwamoto J, Saeki K, Okamoto N. Reliability and validity of the International Physical Activity Questionnaire (IPAQ) in elderly adults: the Fujiwara-kyo Study. *J Epidemiol (Japan).* 2010;21:459–65.
179. van der Ploeg HP, Merom D, Chau JY, Bittman M, Trost SG, Bauman AE. Advances in population surveillance for physical activity and sedentary behavior: reliability and validity of time use surveys. *Am J Epidemiol.* 2010;172:1199–206.
180. van der Ploeg HP, Tudor-Locke C, Marshall AL, Craig C, Hagströmer M, Sjöström M, Bauman A. Reliability and validity of the international physical activity questionnaire for assessing walking. *Res Q Exerc Sport.* 2010;81:97–101.
181. Vandelandotte C, De Bourdeaudhuij I, Philippaerts R, Sjöström M, Sallis J. Reliability and validity of a computerized and Dutch version of the International Physical Activity Questionnaire (IPAQ). *J Phys Act Health.* 2005;2:63–75.
182. Washburn RA, Goldfield SR, Smith KW, McKinlay JB. The validity of self-reported exercise-induced sweating as a measure of physical activity. *Am J Epidemiol.* 1990;132:107–13.
183. Welk GJ, Thompson RW, Galper DI. A temporal validation of scoring algorithms for the 7-day physical activity recall. *Meas Phys Educ Exerc Sci.* 2001;5:123–38.
184. Wendel-Vos GW, Schuit AJ, Saris WH, Kromhout D. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. *J Clin Epidemiol.* 2003;56:1163–9.
185. Wickel EE, Welk GJ, Eisenmann JC. Concurrent validation of the Bouchard Diary with an accelerometry-based monitor. *Med Sci Sports Exerc.* 2006;38:373–9.
186. Wilbur J, Chandler P, Miller AM, Davis GC, Aaronson LS, Mayo K. Measuring adherence to a women's walking program. *West J Nurs Res.* 2001;23:8–32.
187. Yore MM, Ham SA, Ainsworth BE, Kruger J, Reis JP, Kohl H 3rd, Macera CA. Reliability and validity of the instrument used in BRFS5 to assess physical activity. *Med Sci Sports Exerc.* 2007;39:1267–74.
188. Chasan-Taber L, Erickson JB, Nasca PC, Chasan-Taber S, Freedson PS. Validity and reproducibility of a physical activity questionnaire in women. *Med Sci Sports Exerc.* 2002;34:987–92.

189. Kwak L, Kremers SP, Brug J, Van Baak MA. Measuring physical activity in field studies: Comparison of a questionnaire, 24-hour recall and an accelerometer. *Eur J Sport Sci*. 2007;7:193–201.
190. Lagerros YT, Mucci LA, Belloc R, Nyrén O, Bälter O, Bälter KA. Validity and reliability of self-reported total energy expenditure using a novel instrument. *Eur J Epidemiol*. 2006;21:227–36.
191. Richardson MT, Ainsworth BE, Wu H-C, Jacobs DR, Leon AS. Ability of the Atherosclerosis Risk in Communities (ARIC)/Baecke Questionnaire to assess leisure-time physical activity. *Int J Epidemiol*. 1995;24:685–93.
192. Hart TL, Ainsworth BE, Tudor-Locke C. Objective and subjective measures of sedentary behavior and physical activity. *Med Sci Sports Exerc*. 2011;43:449–56.
193. Lee KS, Trost SG. Validity and reliability of the 3-day physical activity recall in Singaporean adolescents. *Res Q Exerc Sport*. 2005;76:101–6.
194. Bassett DR Jr, Ainsworth BE, Swartz AM, Strath SJ, O'Brien WL, King GA. Validity of four motion sensors in measuring moderate intensity physical activity. *Med Sci Sports Exerc*. 2000;32:S471–80.
195. Strath S, Bassett D Jr, Swartz A. Comparison of MTI accelerometer cut-points for predicting time spent in physical activity. *Int J Sports Med*. 2003;24:298–303.
196. Booth ML, Owen N, Bauman AE, Gore CJ. Retest reliability of recall measures of leisure-time physical activity in Australian adults. *Int J Epidemiol*. 1996;25:153–9.
197. Ainsworth BE, Jacobs DR Jr, Leon AS. Validity and reliability of self-reported physical activity status: the Lipid Research Clinics questionnaire. *Med Sci Sports Exerc*. 1993;25:92–8.
198. Ainsworth BE, Sternfeld B, Richardson MT, Jackson K. Evaluation of the kaiser physical activity survey in women. *Med Sci Sports Exerc*. 2000;32:1327–38.
199. Batty D. Reliability of a physical activity questionnaire in middle-aged men. *Public Health*. 2000;114:474–6.
200. Berthouze SE, Minaire PM, Chatard J-C, Boutet C, Castells J, Lacour J-R. A new tool for evaluating energy expenditure: the "QAPSE" development and validation. *Med Sci Sports Exerc*. 1993;25:1405–14.
201. Bharathi A, Sandhya N, Vaz M. The development and characteristics of a physical activity questionnaire for epidemiological studies in urban middle class Indians. *Indian J Med Res*. 2000;111:95–102.
202. Blair SN, Dowda M, Pate RR, Kronenfeld J, Howe HG, Parker G, Blair A, Fridinger F. Reliability of long-term recall of participation in physical activity by middle-aged men and women. *Am J Epidemiol*. 1991;133:266–75.
203. Bouchard C, Tremblay A, Leblanc C, Lortie G, Savard R, Theriault G. A method to assess energy expenditure in children and adults. *Am J Clin Nutr*. 1983;37:461–7.
204. Brown W, Trost S, Bauman A, Mummery K, Owen N. Test-retest reliability of four physical activity measures used in population surveys. *J Sci Med Sport*. 2004;7:205–15.
205. Brownson RC, Eyer AA, King AC, Shyu Y-L, Brown DR, Homan SM. Reliability of information on physical activity and other chronic disease risk factors among US women aged 40 years or older. *Am J Epidemiol*. 1999;149:379–91.
206. Chasan-Taber L, Schmidt MD, Roberts DE, Hosmer D, Markenson G, Freedson PS. Development and validation of a pregnancy physical activity questionnaire. *Med Sci Sports Exerc*. 2004;36:1750–60.
207. Chasan-Taber S, Rimm EB, Stampfer MJ, Spiegelman D, Colditz GA, Giovannucci E, Ascherio A, Willett WC. Reproducibility and validity of a self-administered physical activity questionnaire for male health professionals. *Epidemiology*. 1996;7:81–6.
208. Chinapaw MJ, Sloopmaker SM, Schuit AJ, van Zuidam M, van Mechelen W. Reliability and validity of the Activity Questionnaire for Adults and Adolescents (AQUAA). *BMC Med Res Methodol*. 2009;9:58.
209. Cust AE, Armstrong BK, Smith BJ, Chau J, van der Ploeg HP, Bauman A. Self-reported confidence in recall as a predictor of validity and repeatability of physical activity questionnaire data. *Epidemiology*. 2009;20:433–41.
210. Cust AE, Smith BJ, Chau J, van der Ploeg HP, Friedenreich CM, Armstrong BK, Bauman A. Validity and repeatability of the EPIC physical activity questionnaire: a validation study using accelerometers as an objective measure. *Int J Behav Nutr Phys Act*. 2008;5:33.
211. Cyarto E, Marshall A, Dickinson R, Brown W. Measurement properties of the CHAMPS physical activity questionnaire in a sample of older Australians. *J Sci Med Sport*. 2006;9:319–26.
212. De Abajo S, Larriba R, Marquez S. Validity and reliability of the Yale Physical Activity Survey in Spanish elderly. *J Sports Med Phys Fitness*. 2001;41:479–85.
213. Dinger M, Oman R, Taylor E, Vesely S, Able J. Stability and convergent validity of the Physical Activity Scale for the Elderly (PASE). *J Sports Med Phys Fitness*. 2004;44:186–92.
214. Dipietro L, Caspersen CJ, Ostfeld AM, Nadel ER. A survey for assessing physical activity among older adults. *Med Sci Sports Exerc*. 1993;
215. Dolan SH, Williams DP, Ainsworth BE, Shaw JM. Development and reproducibility of the bone loading history questionnaire. *Med Sci Sports Exerc*. 2006;38:1121–31.
216. Dubbert PM, Vander Weg MW, Kirchner KA, Shaw B. Evaluation of the 7-day physical activity recall in urban and rural men. *Med Sci Sports Exerc*. 2004;36:1646–54.
217. Eagan MS, Lyle RM, George P. A new self-reported comprehensive historical activity questionnaire for young women. *J Phys Act Health*. 2005;2:35.
218. Evenson KR, McGinn AP. Test-retest reliability of adult surveillance measures for physical activity and inactivity. *Am J Prev Med*. 2005;28:470–8.
219. Folsom AR, Jacobs DR, Caspersen CJ, Gomez-Marín O, Knudsen J. Test-retest reliability of the Minnesota leisure time physical activity questionnaire. *J Chronic Dis*. 1986;39:505–11.
220. Friedenreich CM, Courneya KS, Bryant HE. The lifetime total physical activity questionnaire: development and reliability. *Med Sci Sports Exerc*. 1998;30:266–74.
221. Gauthier AP, Larivière M, Young N. Psychometric properties of the IPAQ: a validation study in a sample of northern Franco-Ontarians. *J Phys Act Health*. 2009;6:S54–60.
222. Giles-Corti B, Timperio A, Cutt H, Pikora TJ, Bull FC, Knuijan M, Bulsara M, Van Niel K, Shilton T. Development of a reliable measure of walking within and outside the local neighborhood: RESIDE's Neighborhood Physical Activity Questionnaire. *Prev Med*. 2006;42:455–9.
223. Gill DP, Jones GR, Zou G, Speechley M. The Phone-FIT: a brief physical activity interview for older adults. *J Aging Phys Act*. 2008;16:292–315.
224. Hagiwara A, Ito N, Sawai K, Kazuma K. Validity and reliability of the Physical Activity Scale for the Elderly (PASE) in Japanese elderly people. *Geriatr Gerontol Int*. 2008;8:143–51.
225. Harada ND, Chiu V, King AC, Stewart AL. An evaluation of three self-report physical activity instruments for older adults. *Med Sci Sports Exerc*. 2001;33:962–70.
226. Hopkins W, Wilson N, Worsley F, Russell D. Reliability of the core questionnaire in The Life in New Zealand survey. *N Z J Health Phys*. 1991;24:21–2.
227. Iwai N, Hisamichi S, Hayakawa N, Inaba Y, Nagaoka T, Sugimori H, Seki N, Sakata K, Suzuki K, Tamakoshi A. Validity and reliability of single-item questions about physical activity. *J Epidemiol*. 2001;11:211–8.
228. Jacobs DR Jr, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med Sci Sports Exerc*. 1993;25:81–91.
229. 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;14:98–114.
230. Kurtze N, Rangul V, Hustvedt B-E, Flanders WD. Reliability and validity of self-reported physical activity in the Nord-Trøndelag Health Study (HUNT 2). *Eur J Epidemiol*. 2007;22:379–87.
231. Kurtze N, Rangul V, Hustvedt B-E, Flanders WD. Reliability and validity of self-reported physical activity in the Nord-Trøndelag Health Study—HUNT 1. *Scand J Public Health*. 2008;36:52–61.
232. Lakka TA, Salonen JT. Intra-person variability of various physical activity assessments in the Kuopio Ischaemic Heart Disease Risk Factor Study. *Int J Epidemiol*. 1992;21:467–72.
233. Lee MM, Whittemore AS, Jung DL. Reliability of recalled physical activity, cigarette smoking, and alcohol consumption. *Ann Epidemiol*. 1992;2:705–14.
234. Levin S, Jacobs DR, Ainsworth BE, Richardson MT, Leon AS. Intra-individual variation and estimates of usual physical activity. *Ann Epidemiol*. 1999;9:481–8.
235. Lowther M, Mutrie N, Loughlan C, McFarlane C. Development of a Scottish physical activity questionnaire: a tool for use in physical activity interventions. *Br J Sports Med*. 1999;33:244–9.
236. Matton L, Wijndaele K, Duvigneaud N, Duquet W, Philippaerts R, Thomis M, Lefevre J. Reliability and validity of the Flemish Physical Activity Computerized Questionnaire in adults. *Res Q Exerc Sport*. 2007;78:293–306.
237. Meyer A-M, Evenson KR, Morimoto L, Siscovick D, White E. Test-Retest Reliability of the WHI Physical Activity Questionnaire. *Med Sci Sports Exerc*. 2009;41:530–8.
238. O'Brien-Cousins S. An older adult exercise status inventory: reliability and validity. *Journal of Sport Behavior*. 1996;19:288–306.
239. Orsini N, Belloc R, Bottai M, Pagano M, Wolk A. Reproducibility of the past year and historical self-administered total physical activity questionnaire among older women. *Eur J Epidemiol*. 2007;22:363–8.

240. Pettee GK, McClain JJ, Lee CD, Swan PD, Alvar BA, Mitros MR, Ainsworth BE. Evaluation of physical activity measures used in middle-aged women. *Med Sci Sports Exerc.* 2009;41:1403–12.
241. Philippaerts RM, Lefevre J. Reliability and validity of three physical activity questionnaires in Flemish males. *Am J Epidemiol.* 1998;147:982–90.
242. Pols MA, Peeters P, Kemper HC, Collette H. Repeatability and relative validity of two physical activity questionnaires in elderly women. *Med Sci Sports Exerc.* 1996;28:1020–5.
243. Pols MA, Peeters PH, Bueno-de-Mesquita HB, Ocke MC, Wentink CA, Kemper HC, Collette HJ. Validity and repeatability of a modified Baecke questionnaire on physical activity. *Int J Epidemiol.* 1995;24:381–8.
244. Pols MA, Peeters PH, Ocké MC, Bueno-de-Mesquita HB, Slimani N, Kemper HC, Collette HJ. Relative validity and repeatability of a new questionnaire on physical activity. *Prev Med.* 1997;26:37–43.
245. Rauh MJ, Hovell MF, Hofstetter CR, Sallis JF, Gleghorn A. Reliability and validity of self-reported physical activity in Latinos. *Int J Epidemiol.* 1992;21:966–71.
246. Reed JA, Phillips DA. Relationships between physical activity and the proximity of exercise facilities and home exercise equipment used by undergraduate university students. *J Am Coll Health.* 2005;53:285–90.
247. Reis JP, DuBose KD, Ainsworth BE, Macera CA, Yore MM. Reliability and validity of the occupational physical activity questionnaire. *Med Sci Sports Exerc.* 2005;37:2075–83.
248. Richardson MT, Leon AS, Jacobs DR, Ainsworth BE, Serfass R. Comprehensive evaluation of the Minnesota leisure time physical activity questionnaire. *J Clin Epidemiol.* 1994;47:271–81.
249. Roeykens J, Rogers R, Meeusen R, Magnus L, Borms J, de Meirleir K. Validity and reliability in a Flemish population of the WHO-MONICA Optional Study of Physical Activity Questionnaire. *Med Sci Sports Exerc.* 1998;30:1071–5.
250. Rütten A, Vuillemin A, Ooijendijk W, Schena F, Sjöström M, Stahl T, Vanden Auweele Y, Welshman J, Ziemainz H. Physical activity monitoring in Europe. The European Physical Activity Surveillance System (EUPASS) approach and indicator testing. *Public Health Nutr.* 2003;6:377–84.
251. Sallis JF, Haskell WL, Wood PD, Fortmann SP, Rogers T, Blair SN, Paffenbarger RS. Physical activity assessment methodology in the Five-City Project. *Am J Epidemiol.* 1985;121:91–106.
252. Schuler PB, Richardson MT, Ochoa P, Wang MQ. Accuracy and repeatability of the Yale physical activity survey in assessing physical activity of older adults. *Percept Mot Skills.* 2001;93:163–77.
253. Singh PN, Tonstad S, Abbey DE, Fraser GE. Validity of selected physical activity questions in white Seventh-day Adventists and non-Adventists. *Med Sci Sports Exerc.* 1996;28:1026–37.
254. Smitherman TA, Dubbert PM, Grothe KB, Sung JH, Kendzor DE, Reis JP, Ainsworth BE, Newton RL Jr, Lesniak KT, Taylor HA Jr. Validation of the Jackson Heart Study physical activity survey in African Americans. *J Phys Act Health.* 2009;6:S124–32.
255. Stein AD, Lederman RI, Shea S. The Behavioral Risk Factor Surveillance System questionnaire: its reliability in a statewide sample. *Am J Public Health.* 1993;83:1768–72.
256. Stewart AL, Mills KM, King AC, Haskell WL, Gillis D, Ritter PL. CHAMPS physical activity questionnaire for older adults: outcomes for interventions. *Med Sci Sports Exerc.* 2001;33:1126–41.
257. Trinh OT, Do Nguyen N, van der Ploeg HP, Dibley MJ, Bauman A. Test-retest repeatability and relative validity of the Global Physical Activity Questionnaire in a developing country context. *J Phys Act Health.* 2009;6:S46–53.
258. Voorrips LE, Ravelli AC, Dongelmans A, Deurenberg P, van Staveren WA. A physical activity questionnaire for the elderly. *Med Sci Sports Exerc.* 1991;23:974–9.
259. Vuillemin A, Oppert J-M, Guillemin F, Essermeant L, Fontvieille A-M, Galan P, Kriska AM, Hercberg S. Self-administered questionnaire compared with interview to assess past-year physical activity. *Med Sci Sports Exerc.* 2000;32:1119–24.
260. Wareham NJ, Jakes RW, Rennie KL, Mitchell J, Hennings S, Day NE. Validity and repeatability of the EPIC-Norfolk physical activity questionnaire. *Int J Epidemiol.* 2002;31:168–74.
261. Wareham NJ, Jakes RW, Rennie KL, Schuit J, Mitchell J, Hennings S, Day NE. Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. *Public Health Nutr.* 2003;6:407–13.
262. Washburn R, Cook T, LaPorte R. The objective assessment of physical activity in an occupationally active group. *J Sports Med Phys Fitness.* 1989;29:279–84.
263. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol.* 1993;46:153–62.
264. Weller I, Corey PN. A study of the reliability of the Canada Fitness Survey questionnaire. *Med Sci Sports Exerc.* 1998;30:1530–6.
265. Wilbur J, Holm K, Dan A. A quantitative survey to measure energy expenditure in midlife women. *J Nurs Meas.* 1993;1:29–40.
266. Yasunaga A, Park H, Watanabe E, Togo F, Park S, Shephard RJ, Aoyagi Y. Development and evaluation of the physical activity questionnaire for elderly Japanese: the Nakanajo study. *J Aging Phys Act.* 2007;15:398–411.
267. Yore MM, Ham SA, Ainsworth BE, Macera CA, Jones DA, Kohl HW III. Occupational physical activity: reliability and comparison of activity levels. *J Phys Act Health.* 2005;3.
268. Giles K, Marshall AL. The repeatability and accuracy of CHAMPS as a measure of physical activity in a community sample of older Australian adults. *J Phys Act Health.* 2009;6:221–9.
269. Hayden-Wade HA, Coleman KJ, Sallis JF, Armstrong C. Validation of the telephone and in-person interview versions of the 7-day PAR. *Med Sci Sports Exerc.* 2003;35:801–9.
270. Hamdorf PA, Withers RT, Penhall RK, Haslam MV. Physical training effects on the fitness and habitual activity patterns of elderly women. *Arch Phys Med Rehabil.* 1992;73:603–8.
271. Abel MG, Hannon JC, Sell K, Lillie T, Conlin G, Anderson D. Validation of the Kenz Lifecorder EX and ActiGraph GT1M accelerometers for walking and running in adults. *Appl Physiol Nutr Metab.* 2008;33:1155–64.
272. Allen FR, Ambikairajah E, Lovell NH, Celler BG. Classification of a known sequence of motions and postures from accelerometry data using adapted Gaussian mixture models. *Physiol Meas.* 2006;27:935.
273. Aminian K, Robert P, Buchser E, Rutschmann B, Hayoz D, Depairon M. Physical activity monitoring based on accelerometry: validation and comparison with video observation. *Med Biol Eng Comput.* 1999;37:304–8.
274. Assah FK, Ekelund U, Brage S, Corder K, Wright A, Mbanya JC, Wareham NJ. Predicting Physical Activity Energy Expenditure Using Accelerometry in Adults From Sub-Saharan Africa. *Obesity.* 2009;17:1588–95.
275. Assah FK, Ekelund U, Brage S, Wright A, Mbanya JC, Wareham NJ. Accuracy and validity of a combined heart rate and motion sensor for the measurement of free-living physical activity energy expenditure in adults in Cameroon. *Int J Epidemiol.* 2010;40:112–20.
276. Bonomi AG, Goris A, Yin B, Westerterp KR. Detection of type, duration, and intensity of physical activity using an accelerometer. *Med Sci Sports Exerc.* 2009;41:1770–7.
277. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *J Appl Physiol.* 2009;107:655–61.
278. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. Estimation of Free-Living Energy Expenditure Using a Novel Activity Monitor Designed to Minimize Obtrusiveness. *Obesity.* 2010;18:1845–51.
279. Bussmann JB, Tulen JH, Herel EC, Stam HJ. Quantification of physical activities by means of ambulatory accelerometry: a validation study. *Psychophysiology.* 1998;35:488–96.
280. Carter J, Wilkinson D, Blacker S, Rayson M, Bilzon J, Izard R, Coward A, Wright A, Nevill A, Rennie K. An investigation of a novel three-dimensional activity monitor to predict free-living energy expenditure. *J Sports Sci.* 2008;26:553–61.
281. Choquette S, Chuin A, Lalancette D-A, Brochu M, Dionne IJ. Predicting energy expenditure in elders with the metabolic cost of activities. *Med Sci Sports Exerc.* 2009;41:1915–20.
282. Conger SA, Strath S, Bassett D. Validity and reliability of the FitSense FS-1 speedometer during walking and running. *Int J Sports Med.* 2005;26:208–13.
283. Connolly CP, Coe DP, Kendrick JM, Basset D, Thompson DL. Accuracy of physical activity monitors in pregnant women. *Med Sci Sports Exerc.* 2011;43:1100–5.
284. Dijkstra B, Kamsma Y, Zijlstra W. Detection of gait and postures using a miniaturised triaxial accelerometer-based system: accuracy in community-dwelling older adults. *Arch Phys Med Rehabil.* 2010;39:259–62.
285. Dugas L, Carstens M, Ebersole K, Schoeller D, Durazo-Arvizu R, Lambert E, Luke A. Energy expenditure in young adult urban informal settlement dwellers in South Africa. *Eur J Clin Nutr.* 2009;63:805–7.
286. Ermes M, Parkka J, Mantyjarvi J, Korhonen I. Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions. *IEEE Trans Inf Technol Biomed.* 2008;12:20–6.

287. Eslinger DW, Probert A, Connor GS, Bryan S, Lavolette M, Tremblay MS. Validity of the Actical accelerometer step-count function. *Med Sci Sports Exerc.* 2007;39:1200–4.
288. Godfrey A, Culhane K, Lyons G. Comparison of the performance of the activPAL™ Professional physical activity logger to a discrete accelerometer-based activity monitor. *Med Eng Phys.* 2007;29:930–4.
289. 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;16:201–14.
290. Grant PM, Ryan CG, Tigbe WW, Granat MH. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *Br J Sports Med.* 2006;40:992–7.
291. He Y, Li Y. Physical activity recognition utilizing the built-in kinematic sensors of a smartphone. *Int J Distrib Sens N* 2013, 2013.
292. Jacobi D, Perrin AE, Grosman N, Doré MF, Normand S, Oppert JM, Simon C. Physical Activity-Related Energy Expenditure With the RT3 and TriTrac Accelerometers in Overweight Adults. *Obesity.* 2007;15:950–6.
293. Johannsen DL, Calabro MA, Stewart J, Franke W, Rood JC, Welk GJ. Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Med Sci Sports Exerc.* 2010;42:2134–40.
294. Karantonis DM, Narayanan MR, Mathie M, Lovell NH, Celler BG. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. *IEEE Trans Inf Technol Biomed.* 2006;10:156–67.
295. Koehler K, Braun H, De Marees M, Fusch G, Fusch C, Schaenzer W. Assessing energy expenditure in male endurance athletes: validity of the SenseWear Armband. *Med Sci Sports Exerc.* 2011;43:1328–33.
296. Kumahara H, Schutz Y, Ayabe M, Yoshioka M, Yoshitake Y, Shindo M, Ishii K, Tanaka H. The use of uniaxial accelerometry for the assessment of physical-activity-related energy expenditure: a validation study against whole-body indirect calorimetry. *Br J Nutr.* 2004;91:235–43.
297. Kwon S, Jamal M, Zamba GK, Stumbo P, Samuel I. Validation of a novel physical activity assessment device in morbidly obese females. *J Obes.* 2010;2010.
298. Le Masurier GC, Lee SM, Tudor-Locke C. Motion sensor accuracy under controlled and free-living conditions. *Med Sci Sports Exerc.* 2004;36:905–10.
299. Le Masurier GC, Tudor-Locke C. Comparison of pedometer and accelerometer accuracy under controlled conditions. *Med Sci Sports Exerc.* 2003;35:867–71.
300. Leenders NY, Sherman WM, Nagaraja HN. Energy expenditure estimated by accelerometry and doubly labeled water: do they agree? *Med Sci Sports Exerc.* 2006;38:2165–72.
301. Levine JA, McCrady SK, Lanningham-Foster LM, Kane PH, Foster RC, Manohar CU. The role of free-living daily walking in human weight gain and obesity. *Diabetes.* 2008;57:548–54.
302. Liu S-H, Chang Y-J. Using accelerometers for physical actions recognition by a neural fuzzy network. *Telemedicine and e-Health.* 2009;15:867–76.
303. Löf M. Physical activity pattern and activity energy expenditure in healthy pregnant and non-pregnant Swedish women. *Eur J Clin Nutr.* 2011;65:1295–301.
304. Lof M, Hannestad U, Forsum E. Comparison of commonly used procedures, including the doubly-labelled water technique, in the estimation of total energy expenditure of women with special reference to the significance of body fatness. *Br J Nutr.* 2003;90:961–8.
305. Maddison R, Jiang Y, Vander Hoorn S, Ni Mhurchu C, Lawes CM, Rodgers A, Rush E. Estimating energy expenditure with the RT3 triaxial accelerometer. *Res Q Exerc Sport.* 2009;80:249–56.
306. Maddocks M, Petrou A, Skipper L, Wilcock A. Validity of three accelerometers during treadmill walking and motor vehicle travel. *Br J Sports Med.* 2010;44:606–8.
307. Marsh AP, Vance RM, Frederick TL, Hesselmann SA, Rejeski WJ. Objective assessment of activity in older adults at risk for mobility disability. *Med Sci Sports Exerc.* 2007;39:1020–6.
308. Mathie M, Celler BG, Lovell NH, Coster A. Classification of basic daily movements using a triaxial accelerometer. *Med Biol Eng Comput.* 2004;42:679–87.
309. Mathie M, Coster A, Lovell N, Celler B. Detection of daily physical activities using a triaxial accelerometer. *Med Biol Eng Comput.* 2003;41:296–301.
310. Najafi B, Aminian K, Paraschiv-Ionescu A, Loew F, Büla CJ, Robert P. Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly. *IEEE Trans Biomed Eng.* 2003;50:711–23.
311. Plasqui G, Joosen AM, Kester AD, Goris AH, Westerterp KR. Measuring Free-Living Energy Expenditure and Physical Activity with Triaxial Accelerometry. *Obes Res.* 2005;13:1363–9.
312. Preece SJ, Goulermas JY, Kenney LP, Howard D. A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. *IEEE Trans Biomed Eng.* 2009;56:871–9.
313. Rafamantanantsoa HH, Ebine N, Yoshioka M, Yoshitake Y, Tanaka H, Saitoh S, Jones PJH. The role of exercise physical activity in varying the total energy expenditure in healthy Japanese men 30 to 69 years of age. *J Nutr Sci Vitaminol (Tokyo).* 2003;49:120–4.
314. Rowlands AV, Stone MR, Eston RG. Influence of speed and step frequency during walking and running on motion sensor output. *Med Sci Sports Exerc.* 2007;39:716–27.
315. Ryan CG, Grant PM, Tigbe WW, Granat MH. The validity and reliability of a novel activity monitor as a measure of walking. *Br J Sports Med.* 2006;40:779–84.
316. Staudenmayer J, Pober D, Crouter S, Bassett D, Freedson P. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *J Appl Physiol.* 2009;107:1300–7.
317. St-Onge M, Mignault D, Allison DB, Rabasa-Lhoret R. Evaluation of a portable device to measure daily energy expenditure in free-living adults. *Am J Clin Nutr.* 2007;85:742–9.
318. Storti KL, Pettee KK, Brach JS, Talkowski JB, Richardson CR, Kriska AM. Gait speed and step-count monitor accuracy in community-dwelling older adults. *Med Sci Sports Exerc.* 2008;40:59–64.
319. Wu W, Dasgupta S, Ramirez EE, Peterson C, Norman GJ. Classification accuracies of physical activities using smartphone motion sensors. *J Med Internet Res.* 2012;14:e130.
320. Yang J-Y, Wang J-S, Chen Y-P. Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers. *Pattern Recogn Lett.* 2008;29:2213–20.
321. Zhang K, Werner P, Sun M, Pi-Sunyer FX, Boozer CN. Measurement of human daily physical activity. *Obes Res.* 2003;11:33–40.
322. Lee MH, Kim J, Jee SH, Yoo SK. Integrated solution for physical activity monitoring based on mobile phone and PC. *Healthc Inform Res.* 2011;17:76–86.
323. Mackey DC, Manini TM, Schoeller DA, Koster A, Glynn NW, Goodpaster BH, Satterfield S, Newman AB, Harris TB, Cummings SR. Validation of an armband to measure daily energy expenditure in older adults. *J Gerontol A Biol Sci Med Sci.* 2011;66A:1108–13.
324. Anderson I, Maitland J, Sherwood S, Barkhuus L, Chalmers M, Hall M, Brown B, Muller H. Shakra: tracking and sharing daily activity levels with unaugmented mobile phones. *Mobile Networks and Applications.* 2007;12:185–99.
325. Atallah L, Leong J, Lo B, Yang G-Z. Energy expenditure prediction using a miniaturized ear-worn sensor. *Med Sci Sports Exerc.* 2011;43:1369–77.
326. Ayen TG, Montoye HJ. Estimation of energy expenditure with a simulated three-dimensional accelerometer. *Journal of Ambulatory Monitoring.* 1988;1:293–301.
327. Balogun JA, Amusa LO, Onyewadume IU. Factors affecting Caltrac® and Calcount® accelerometer output. *Phys Ther.* 1988;68:1500–4.
328. Balogun JA, Martin DA, Clendenin MA. Calorimetric validation of the Caltrac® accelerometer during level walking. *Phys Ther.* 1989;69:501–9.
329. Bao L, Intille SS. Activity recognition from user-annotated acceleration data. *Pervasive Computing.* 2004;3001:1–17.
330. Barreira TV, Brouillette RM, Foil HC, Keller JN, Tudor-Locke C. Comparison of Older Adults' Steps per Day Using an NL-1000 Pedometer and Two GT3X+ Accelerometer Filters. *J Aging Phys Act.* 2013;21:402–16.
331. Berntsen S, Hageberg R, Aandstad A, Mowinckel P, Anderssen SA, Carlsen K, Andersen LB. Validity of physical activity monitors in adults participating in free-living activities. *Br J Sports Med.* 2010;44:657–64.
332. Berntsen S, Stafne SN, Mørkved S. Physical activity monitor for recording energy expenditure in pregnancy. *Acta Obstet Gynecol Scand.* 2011;90:903–7.
333. Bouten CV, Koekkoek KT, Verduin M, Kodde R, Janssen JD. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Trans Biomed Eng.* 1997;44:136–47.
334. Bouten CV, Sauren AA, Verduin M, Janssen J. Effects of placement and orientation of body-fixed accelerometers on the assessment of energy expenditure during walking. *Med Biol Eng Comput.* 1997;35:50–6.
335. Bouten CV, Westerterp KR, Verduin M, Janssen J. Assessment of energy expenditure for physical activity using a triaxial accelerometer. *Med Sci Sports Exerc.* 1994;26:1516–23.
336. Brage S, Brage N, Franks P, Ekelund U, Wareham N. Reliability and validity of the combined heart rate and movement sensor Actiheart. *Eur J Clin Nutr.* 2005;59:561–70.

337. Brage S, Wedderkopp N, Franks PW, Andersen LB, Froberg K. Reexamination of validity and reliability of the CSA monitor in walking and running. *Med Sci Sports Exerc.* 2003;35:1447–54.
338. Brooks AG, Gunn SM, Withers RT, Gore CJ, Plummer JL. Predicting walking METs and energy expenditure from speed or accelerometry. *Med Sci Sports Exerc.* 2005;37:1216–23.
339. Brugniaux JV, Niva A, Pulkkinen I, Laukkanen RM, Richalet J-P, Pichon AP. Polar Activity Watch 200: a new device to accurately assess energy expenditure. *Br J Sports Med.* 2010;44:245–9.
340. Cain KL, Conway TL, Adams MA, Husak LE, Sallis JF. Comparison of older and newer generations of ActiGraph accelerometers with the normal filter and the low frequency extension. *Int J Behav Nutr Phys Act.* 2013;10:51.
341. Campbell KL, Crocker P, McKenzie DC. Field evaluation of energy expenditure in women using Tritrac accelerometers. *Med Sci Sports Exerc.* 2002;34:1667–74.
342. Chen KY, Acra SA, Majchrzak K, Donahue CL, Baker L, Clemens L, Sun M, Buchowski MS. Predicting energy expenditure of physical activity using hip- and wrist-worn accelerometers. *Diabetes Technology and Therapeutics.* 2003;5:1023–33.
343. Chen KY, Sun M. Improving energy expenditure estimation by using a triaxial accelerometer. *J Appl Physiol.* 1997;83:2112–22.
344. Copeland JL, Eslinger DW. Accelerometer assessment of physical activity in active, healthy older adults. *J Aging Phys Act.* 2009;17:17–30.
345. Crouter SE, Bassett DR. A new 2-regression model for the Actical accelerometer. *Br J Sports Med.* 2008;42:217–24.
346. Crouter SE, Churilla JR, Bassett DR. Accuracy of the Actiheart for the assessment of energy expenditure in adults. *Eur J Clin Nutr.* 2008;62:704–11.
347. Crouter SE, Churilla JR, Bassett DR Jr. Estimating energy expenditure using accelerometers. *Eur J Appl Physiol.* 2006;98:601–12.
348. Crouter SE, Clowers KG, Bassett DR. A novel method for using accelerometer data to predict energy expenditure. *J Appl Physiol.* 2006;100:1324–31.
349. Crouter SE, DellaValle DM, Haas JD, Frongillo EA, Bassett DR. Validity of actiGraph 2-regression model and Matthews and NHANES and cut-points for assessing free-living physical activity. *J Phys Act Health.* 2013;10:504–14.
350. Crouter SE, DellaValle DM, Horton M, Haas JD, Frongillo EA, Bassett DR Jr. Validity of the Actical for estimating free-living physical activity. *Eur J Appl Physiol.* 2011;111:1381–9.
351. DeVoe D, Gotshall R, McArthur T. Comparison of the RT3 Research Tracker and Tritrac R3D accelerometers. *Percept Mot Skills.* 2003;97:510–8.
352. Donaire-Gonzalez D, de Nazelle A, Seto E, Mendez M, Nieuwenhuijsen MJ, Jerrett M. Comparison of physical activity measures using mobile phone-based CalFit and Actigraph. *J Med Internet Res.* 2013;15:111.
353. Drenowatz C, Eisenmann JC. Validation of the SenseWear Armband at high intensity exercise. *Eur J Appl Physiol.* 2011;111:883–7.
354. Duncan GE, Lester J, Migotsky S, Goh J, Higgins L, Borriello G. Accuracy of a novel multi-sensor board for measuring physical activity and energy expenditure. *Eur J Appl Physiol.* 2011;111:2025–32.
355. Eslinger DW, Rowlands AV, Hurst TL, Catt M, Murray P, Eston RG. Validation of the GENEA Accelerometer. *Med Sci Sports Exerc.* 2011;43:1085–93.
356. Fehling PC, Smith DL, Warner SE, Dalsky GP. Comparison of accelerometers with oxygen consumption in older adults during exercise. *Med Sci Sports Exerc.* 1999;31:171–5.
357. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc.* 1998;30:777–81.
358. Fruin ML, Rankin JW. Validity of a multi-sensor armband in estimating rest and exercise energy expenditure. *Med Sci Sports Exerc.* 2004;36:1063–9.
359. Fudge BW, Wilson J, Easton C, Irwin L, Clark J, Haddow O, Kayser B, Pitsiladis YP. Estimation of oxygen uptake during fast running using accelerometry and heart rate. *Med Sci Sports Exerc.* 2007;39:192–8.
360. Hageman PA, Norman JF, Pfefferkorn KL, Reiss NJ, Riesberg KA. Comparison of two physical activity monitors during a 1-mile walking field test. *J Exerc Physiol Online.* 2004;7:102–10.
361. Ham SA, Reis JP, Strath SJ, Dubose KD, Ainsworth BE. Discrepancies between methods of identifying objectively determined physical activity. *Med Sci Sports Exerc.* 2007;39:52–8.
362. Harrington DM, Welk GJ, Donnelly AE. Validation of MET estimates and step measurement using the ActivPAL physical activity logger. *J Sports Sci.* 2011; 29:627–33.
363. Haymes EM, Byrnes WC. Walking and running energy expenditure estimated by Caltrac and indirect calorimetry. *Med Sci Sports Exerc.* 1993;25:1365–9.
364. vanHees VT, Lummel RC, Westerterp KR. Estimating Activity-related Energy Expenditure Under Sedentary Conditions Using a Tri-axial Seismic Accelerometer. *Obesity.* 2009;17:1287–92.
365. Hendelman D, Miller K, Baggett C, Debold E, Freedson P. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Med Sci Sports Exerc.* 2000;32:S442–9.
366. Herman Hansen B, Bortnes I, Hildebrand M, Holme I, Kolle E, Anderssen SA. Validity of the ActiGraph GT1M during walking and cycling. *J Sports Sci.* 2014;32:510–6.
367. Howe CA, Staudenmayer JW, Freedson PS. Accelerometer prediction of energy expenditure: vector magnitude versus vertical axis. *Med Sci Sports Exerc.* 2009;41:2199–206.
368. Hustvedt B-E, Christophersen A, Johnsen LR, Tomten H, McNeill G, Haggarty P, Løvø A. Description and validation of the ActiReg®: a novel instrument to measure physical activity and energy expenditure. *Br J Nutr.* 2004;92:1001–8.
369. Jakicic JM, Marcus M, Gallagher KI, Randall C, Thomas E, Goss FL, Robertson RJ. Evaluation of the SenseWear Pro Armband to assess energy expenditure during exercise. *Med Sci Sports Exerc.* 2004;36:897–904.
370. Jakicic JM, Winters C, Lagally K, Ho J, Robertson RJ, Wing RR. The accuracy of the TriTrac-R3D accelerometer to estimate energy expenditure. *Med Sci Sports Exerc.* 1999;31:747–54.
371. Johansson HP, Rossander-Hulthén L, Slinde F, Ekblom B. Accelerometry combined with heart rate telemetry in the assessment of total energy expenditure. *Br J Nutr.* 2006;95:631–9.
372. John D, Tyo B, Bassett DR. Comparison of four ActiGraph accelerometers during walking and running. *Med Sci Sports Exerc.* 2010;42:368–74.
373. Kane N, Simmons M, John D, Thompson D, Bassett D, Bassett D. Validity of the Nike+ device during walking and running. *Int J Sports Med.* 2010;31:101–5.
374. King GA, Torres N, Potter C, Brooks TJ, Coleman KJ. Comparison of activity monitors to estimate energy cost of treadmill exercise. *Med Sci Sports Exerc.* 2004;36:1244–51.
375. Kuffel EE, Crouter SE, Haas JD, Frongillo EA, Bassett DR Jr. Validity of estimating minute-by-minute energy expenditure of continuous walking bouts by accelerometry. *Int J Behav Nutr Phys Act.* 2011;8:92.
376. Leenders NY, Nelson TE, Sherman WM. Ability of different physical activity monitors to detect movement during treadmill walking. *Int J Sports Med.* 2003;24:43–50.
377. Levine J, Melanson EL, Westerterp KR, Hill JO. Measurement of the components of nonexercise activity thermogenesis. *Am J Physiol Endocrinol Metab.* 2001;281:E670–5.
378. Levine JA. An accelerometer-based earpiece to monitor and quantify physical activity. *J Phys Act Health.* 2009;6:781–9.
379. Lyden K, Kozey SL, Staudenmayer JW, Freedson PS. A comprehensive evaluation of commonly used accelerometer energy expenditure and MET prediction equations. *Eur J Appl Physiol.* 2011;111:187–201.
380. Maliszewski AF, Freedson PS, Ebbeling CJ, Crusemeyer J, Kastango KB. Validity of the Caltrac accelerometer in estimating energy expenditure and activity in children and adults. *Pediatr Exerc Sci.* 1991;3:141–51.
381. Matthews CE, Freedson PS, Hebert JR, Stanek E 3rd, Merriam PA, Ockene IS. Comparing physical activity assessment methods in the Seasonal Variation of Blood Cholesterol Study. *Med Sci Sports Exerc.* 2000;32:976–84.
382. Melanson EL Jr, Freedson PS. Validity of the Computer Science and Applications, Inc.(CSA) activity monitor. *Med Sci Sports Exerc.* 1995;27:934–40.
383. Midorikawa T, Tanaka S, Kaneko K, Koizumi K, Ishikawa-Takata K, Futami J, Tabata I. Evaluation of Low-Intensity Physical Activity by Triaxial Accelerometry. *Obesity.* 2007;15:3031–8.
384. Miller NE, Strath SJ, Swartz AM, Cashin SE. Estimating absolute and relative physical activity intensity across age via accelerometry in adults. *J Aging Phys Act.* 2010;18:158.
385. Nichols JF, Morgan CG, Sarkin JA, Sallis JF, Calfas KJ. Validity, reliability, and calibration of the Tritrac accelerometer as a measure of physical activity. *Med Sci Sports Exerc.* 1999;31:908–12.
386. Ohkawara K, Oshima Y, Hikiyama Y, Ishikawa-Takata K, Tabata I, Tanaka S. Real-time estimation of daily physical activity intensity by a triaxial accelerometer and a gravity-removal classification algorithm. *Br J Nutr.* 2011; 105:1681–91.
387. Pambianco G, Wing RR, Robertson R. Accuracy and reliability of the Caltrac accelerometer for estimating energy expenditure. *Med Sci Sports Exerc.* 1990;22:858–62.

388. Patterson SM, Krantz DS, Montgomery LC, Deuster PA, Hedges SM, Nebel LE. Automated physical activity monitoring: Validation and comparison with physiological and self-report measures. *Psychophysiology*. 1993;30:296–305.
389. Paul DR, Kramer M, Moshfegh AJ, Baer DJ, Rumpler WV. Comparison of two different physical activity monitors. *BMC Med Res Methodol*. 2007;7:26.
390. Pober DM, Staudenmayer J, Raphael C, Freedson PS. Development of novel techniques to classify physical activity mode using accelerometers. *Med Sci Sports Exerc*. 2006;38:1626–34.
391. Pomeroy J, Brage S, Curtis JM, Swan PD, Knowler WC, Franks PW. Between-monitor differences in step counts are related to body size: implications for objective physical activity measurement. *PLoS One*. 2011;6:e18942.
392. Ried-Larsen M, Brønd JC, Brage S, Hansen BH, Grydeland M, Andersen LB, Møller NC. Mechanical and free living comparisons of four generations of the Actigraph activity monitor. *Int J Behav Nutr Phys Act*. 2012;9:113.
393. Rothney MP, Neumann M, Beziat A, Chen KY. An artificial neural network model of energy expenditure using nonintegrated acceleration signals. *J Appl Physiol*. 2007;103:1419–27.
394. Rothney MP, Schaefer EV, Neumann MM, Choi L, Chen KY. Validity of physical activity intensity predictions by ActiGraph, Actical, and RT3 accelerometers. *Obesity*. 2008;16:1946–52.
395. Rowlands AV, Thomas P, Eston RG, Topping R. Validation of the RT3 triaxial accelerometer for the assessment of physical activity. *Med Sci Sports Exerc*. 2004;36:518–24.
396. Sherman W, Morris D, Kirby T, Petosa R, Smith B, Frid D, Leenders N. Evaluation of a commercial accelerometer (Tritrac-R3 D) to measure energy expenditure during ambulation. *Int J Sports Med*. 1998;19:43–7.
397. Sloomaker S, Paw CA, Schuit A, Van Mechelen W, Koppes L. Concurrent validity of the PAM accelerometer relative to the MTI Actigraph using oxygen consumption as a reference. *Scand J Med Sci Sports*. 2009;19:36–43.
398. Spierer DK, Hagins M, Rundle A, Pappas E. A comparison of energy expenditure estimates from the Actiheart and Actical physical activity monitors during low intensity activities, walking, and jogging. *Eur J Appl Physiol*. 2011;111:659–67.
399. Stone MR, Eslinger DW, Tremblay MS. Comparative validity assessment of five activity monitors: Does being a child matter? *Pediatr Exerc Sci*. 2007;19:291–309.
400. Strath SJ, Bassett DR Jr, Swartz A, Thompson DL. Simultaneous heart rate-motion sensor technique to estimate energy expenditure. *Med Sci Sports Exerc*. 2001;33:2118–23.
401. Swan PD, Byrnes WC, Haymes EM. Energy expenditure estimates of the Caltrac accelerometer for running, race walking, and stepping. *Br J Sports Med*. 1997;31:235–9.
402. Swartz AM, Strath SJ, Bassett DR Jr, O'Brien WL, King GA, Ainsworth BE. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Med Sci Sports Exerc*. 2000;32:S450–6.
403. Terrier P, Aminian K, Schutz Y. Can accelerometry accurately predict the energy cost of uphill/downhill walking? *Ergonomics*. 2001;44:48–62.
404. Thompson D, Batterham AM, Bock S, Robson C, Stokes K. Assessment of low-to-moderate intensity physical activity thermogenesis in young adults using synchronized heart rate and accelerometry with branched-equation modeling. *The Journal of Nutrition*. 2006;136:1037–42.
405. Wanner M, Martin BW, Meier F, Probst-Hensch N, Kriemler S. Effects of filter choice in GT3X accelerometer assessments of free-living activity. *Med Sci Sports Exerc*. 2013;45:170–7.
406. Welk GJ, Almeida J, Morss G. Laboratory calibration and validation of the Biotrainer and Actitrac activity monitors. *Med Sci Sports Exerc*. 2003;35:1057–64.
407. Welk GJ, Blair SN, Wood K, Jones S, Thompson RW. A comparative evaluation of three accelerometer-based physical activity monitors. *Med Sci Sports Exerc*. 2000;32:S489–97.
408. Welk GJ, McClain JJ, Eisenmann JC, Wickel EE. Field validation of the MTI Actigraph and BodyMedia armband monitor using the IDEEA monitor. *Obesity*. 2007;15:918–28.
409. Yngve A, Nilsson A, Sjöström M, Ekelund U. Effect of monitor placement and of activity setting on the MTI accelerometer output. *Med Sci Sports Exerc*. 2003;35:320–6.
410. Brage S, Brage N, Ekelund U, Luan J, Franks PW, Froberg K, Wareham NJ. Effect of combined movement and heart rate monitor placement on physical activity estimates during treadmill locomotion and free-living. *Eur J Appl Physiol*. 2006;96:517–24.
411. Horner FE, Rayson M, Bilzon JL. Reliability and validity of the 3DNX accelerometer during mechanical and human treadmill exercise testing. *Int J Obes*. 2011;35:S88–97.
412. Kumahara H, Tanaka H, Schutz Y. Daily physical activity assessment: what is the importance of upper limb movements vs whole body movements? *Int J Obes*. 2004;28:1105–10.
413. McClain JJ, Sisson SB, Tudor-Locke C. Actigraph accelerometer interinstrument reliability during free-living in adults. *Med Sci Sports Exerc*. 2007;39:1509–14.
414. Meijer GA, Westerterp KR, Verhoeven FM, Koper HB, ten Hoor F. Methods to assess physical activity with special reference to motion sensors and accelerometers. *IEEE Transactions on Biomedical Engineering*. 1991;38:221–9.
415. Nichols JF, Morgan CG, Chabot LE, Sallis JF, Calfas KJ. Assessment of physical activity with the Computer Science and Applications, Inc., accelerometer: laboratory versus field validation. *Res Q Exerc Sport*. 2000;71:36–43.
416. Powell SM, Rowlands AV. Intermonitor variability of the RT3 accelerometer during typical physical activities. *Med Sci Sports Exerc*. 2004;36:324–30.
417. Reneman M, Helmus M. Interinstrument reliability of the RT3 accelerometer. *Int J Rehabil Res*. 2010;33:178–9.
418. Vanhelst J, Baquet G, Gottrand F, Béghin L. Comparative interinstrument reliability of uniaxial and triaxial accelerometers in free-living conditions 1, 2. *Percept Mot Skills*. 2012;114:584–94.
419. Brandes M, Rosenbaum D. Correlations between the step activity monitor and the DynaPort ADL-monitor. *Clinical Biomechanics*. 2004;19:91–4.
420. Dahlgren G, Carlsson D, Moorhead A, Häger-Ross C, McDonough SM. Test-retest reliability of step counts with the ActivPAL™ device in common daily activities. *Gait Posture*. 2010;32:386–90.
421. Gao KL, Tsang WW. Use of accelerometry to quantify the physical activity level of the elderly. *H K Physio J*. 2008;26:18–23.
422. Harris TJ, Owen CG, Victor CR, Adams R, Ekelund U, Cook DG. A comparison of questionnaire, accelerometer, and pedometer: measures in older people. *Med Sci Sports Exerc*. 2009;41:1392–402.
423. Hart TL, Swartz AM, Cashin SE, Strath SJ. How many days of monitoring predict physical activity and sedentary behaviour in older adults. *Int J Behav Nutr Phys Act*. 2011;16:62.
424. Matthews CE, Ainsworth BE, Thompson RW, Bassett DR Jr. Sources of variance in daily physical activity levels as measured by an accelerometer. *Med Sci Sports Exerc*. 2002;34:1376–81.
425. Montoye HJ, Washburn R, Servais S, Ertl A, Webster JG, Nagle FJ. Estimation of energy expenditure by a portable accelerometer. *Med Sci Sports Exerc*. 1983;15:403–7.
426. Nichols J, Patterson P, Early T. A validation of a physical activity monitor for young and older adults. *Can J Sport Sci*. 1992;17:299–303.
427. Rowe DA, Kembler CD, Robinson TS, Mahar MT. Daily walking in older adults: day-to-day variability and criterion-referenced validity of total daily step counts. *J Phys Act Health*. 2007;4:435–47.
428. Sirard JR, Forsyth A, Michael Oakes J, Schmitz KH. Accelerometer test-retest reliability by data processing algorithms: results from the Twin Cities Walking Study. *J Phys Act Health*. 2011;8:668–74.
429. Sugimoto A, Hara Y, Findley T, Yoncmoto K. A useful method for measuring daily physical activity by a three-direction monitor. *Scand J Rehabil Med*. 1997;29:37–42.
430. Welk GJ, Schaben JA, Morrow JR Jr. Reliability of accelerometry-based activity monitors: a generalizability study. *Med Sci Sports Exerc*. 2004;36:1637–45.
431. Baranowski T, Masse LC, Ragan B, Welk G. How many days was that? We're still not sure, but we're asking the question better! *Med Sci Sports Exerc*. 2008;40:S544.
432. Ayabe M, Ishii K, Takayama K, Aoki J, Tanaka H. Comparison of interdevice measurement difference of pedometers in younger and older adults. *Br J Sports Med*. 2010;44:95–9.
433. Bergman RJ, Bassett D, Klein DA. Validity of 2 devices for measuring steps taken by older adults in assisted-living facilities. *J Phys Act Health*. 2008;5:5166–75.
434. Clemes SA, O'Connell S, Rogan LM, Griffiths PL. Evaluation of a commercially available pedometer used to promote physical activity as part of a national programme. *Br J Sports Med*. 2010;44:1178–83.
435. Crouter SE, Schneider PL, Bassett DR Jr. Spring-levered versus piezo-electric pedometer accuracy in overweight and obese adults. *Med Sci Sports Exerc*. 2005;37:1673–9.
436. Cyarto EV, Myers A, Tudor-Locke C. Pedometer accuracy in nursing home and community-dwelling older adults. *Med Sci Sports Exerc*. 2004;36:205–9.
437. Foster RC, Lanningham-Foster LM, Manohar C, McCrady SK, Nysse LJ, Kaufman KR, Padgett DJ, Levine JA. Precision and accuracy of an ankle-worn

- accelerometer-based pedometer in step counting and energy expenditure. *Prev Med.* 2005;41:778–83.
438. Hasson R, Haller J, Pober D, Staudenmayer J, Freedson P. Validity of the Omron HJ-112 pedometer during treadmill walking. *Med Sci Sports Exerc.* 2009;41:805–9.
439. Karabulut M, Crouter SE, Bassett DR Jr. Comparison of two waist-mounted and two ankle-mounted electronic pedometers. *Eur J Appl Physiol.* 2005;95:335–43.
440. Melanson EL, Knoll JR, Bell ML, Donahoo WT, Hill J, Nysse LJ, Lanningham-Foster L, Peters JC, Levine JA. Commercially available pedometers: considerations for accurate step counting. *Prev Med.* 2004;39:361–8.
441. Park J, Ishikawa-Takata K, Tanaka S, Mekata Y, Tabata I. Effects of walking speed and step frequency on estimation of physical activity using accelerometers. *J Physiol Anthropol.* 2011;30:119–27.
442. Shepherd EF, Toloza E, McClung CD, Schmalzried TP. Step activity monitor: increased accuracy in quantifying ambulatory activity. *J Orthop Res.* 1999;17:703–8.
443. Swartz A, Bassett D Jr, Moore J, Thompson D, Strath S. Effects of body mass index on the accuracy of an electronic pedometer. *Int J Sports Med.* 2003;24:588–92.
444. Tudor-Locke C, Sisson SB, Lee SM, Craig CL, Plotnikoff RC, Bauman A. Evaluation of quality of commercial pedometers. *Can J Public Health.* 2006;97:S10–5.
445. Welk GJ, Differding JA, Thompson RW, Blair SN, Dziura J, Hart P. The utility of the Digi-walker step counter to assess daily physical activity patterns. *Med Sci Sports Exerc.* 2000;32:S481–8.
446. Schneider PL, Crouter SE, Lukajic O, Bassett DR. Accuracy and reliability of 10 pedometers for measuring steps over a 400-m walk. *Med Sci Sports Exerc.* 2003;35:1779–84.
447. Holbrook E, Barreira T, Kang M. Validity and reliability of Omron pedometers for prescribed and self-paced walking. *Med Sci Sports Exerc.* 2009;41:669–3.
448. Abel MG, Hannon JC, Eisenman PA, Ransdell LB, Pett M, Williams DP. Waist circumference, pedometer placement, and step-counting accuracy in youth. *Res Q Exerc Sport.* 2009;80:434–44.
449. Crouter SE, Schneider PL, Karabulut M, Bassett DR Jr. Validity of 10 electronic pedometers for measuring steps, distance, and energy cost. *Med Sci Sports Exerc.* 2003;35:1455–60.
450. Resnick B, Nahm E-S, Orwig D, Zimmerman SS, Magaziner J. Measurement of activity in older adults: reliability and validity of the Step Activity Monitor. *J Nurs Meas.* 2001;9:275–90.
451. Bassett DR Jr, Ainsworth BE, Leggett SR, Mathien CA, Main JA, Hunter DC, Duncan GE. Accuracy of five electronic pedometers for measuring distance walked. *Med Sci Sports Exerc.* 1996;28:1071–7.
452. McClain JJ, Hart TL, Getz RS, Tudor-Locke C. Convergent validity of 3 low cost motion sensors with the ActiGraph accelerometer. *J Phys Act Health.* 2010;7:662–70.
453. Tharion WJ, Yokota M, Buller MJ, DeLany JP, Hoyt RW. Total energy expenditure estimated using a foot-contact pedometer. *Am J Case Rep.* 2004;10:504–9.
454. Hoyt RW, Buller MJ, Santee WR, Yokota M, Weyand PG, Delany JP. Total energy expenditure estimated using foot-ground contact pedometry. *Diabetes Technology and Therapeutics.* 2004;6:71–81.
455. Barreira TV, Tudor-Locke C, Champagne CM, Broyles ST, Johnson WD, Katzmarzyk PT. Comparison of GT3X accelerometer and YAMAX pedometer steps/day in a free-living sample of overweight and obese adults. *J Phys Act Health.* 2013;10:263–70.
456. Behrens TK, Dinger MK. Comparisons of accelerometer and pedometer determined steps in free living samples. *J Phys Act Health.* 2011;8:390.
457. McClain JJ, Craig CL, Sisson SB, Tudor-Locke C. Comparison of Lifecorder EX and ActiGraph accelerometers under free-living conditions. *Appl Physiol Nutr Metab.* 2007;32:753–61.
458. Schneider PL, Crouter SE, Bassett DR. Pedometer measures of free-living physical activity: comparison of 13 models. *Med Sci Sports Exerc.* 2004;36:331–5.
459. Tudor-Locke C, Ainsworth BE, Thompson RW, Matthews CE. Comparison of pedometer and accelerometer measures of free-living physical activity. *Med Sci Sports Exerc.* 2002;34:2045–51.
460. Felton GM, Tudor-Locke C, Burkett L. Reliability of pedometer-determined free-living physical activity data in college women. *Res Q Exerc Sport.* 2006;77:304–8.
461. Strycker LA, Duncan SC, Chaumeton NR, Duncan TE, Toobert DJ. Reliability of pedometer data in samples of youth and older women. *Int J Behav Nutr Phys Act.* 2007;4:4.
462. Kang M, Bassett DR, Barreira TV, Tudor-Locke C, Ainsworth B, Reis JP, Strath S, Swartz A. How many days are enough? A study of 365 days of pedometer monitoring. *Res Q Exerc Sport.* 2009;80:445–53.
463. Tudor-Locke C, Burkett L, Reis J, Ainsworth B, Macera C, Wilson D. How many days of pedometer monitoring predict weekly physical activity in adults? *Prev Med.* 2005;40:293–8.
464. Tudor-Locke C. A preliminary study to determine instrument responsiveness to change with a walking program: physical activity logs versus pedometers. *Res Q Exerc Sport.* 2001;72:288–92.
465. Kashiwazaki H. Heart rate monitoring as a field method for estimating energy expenditure as evaluated by the doubly labeled water method. *J Nutr Sci Vitaminol (Tokyo).* 1999;45:79–94.
466. Livingstone M, Prentice AM, Coward W, Ceesay SM, Strain JJ, McKenna PG, Nevin GB, Barker ME, Hickey R. Simultaneous measurement of free-living energy expenditure by the doubly labeled water method and heart-rate monitoring. *Am J Clin Nutr.* 1990;52:59–65.
467. Ceesay SM, Prentice AM, Day KC, Murgatroyd PR, Goldberg GR, Scott W, Spurr G. The use of heart rate monitoring in the estimation of energy expenditure: a validation study using indirect whole-body calorimetry. *Br J Nutr.* 1989;61:175–86.
468. Leonard WR, Katzmarzyk PT, Stephen MA, Ross A. Comparison of the heart rate-monitoring and factorial methods: assessment of energy expenditure in highland and coastal Ecuadoreans. *Am J Clin Nutr.* 1995;61:1146–52.
469. Spurr G, Prentice A, Murgatroyd P, Goldberg G, Reina J, Christman N. Energy expenditure from minute-by-minute heart-rate recording: comparison with indirect calorimetry. *Am J Clin Nutr.* 1988;48:552–9.
470. Strath SJ, Swartz AM, Bassett DR Jr, O'Brien WL, King GA, Ainsworth BE. Evaluation of heart rate as a method for assessing moderate intensity physical activity. *Med Sci Sports Exerc.* 2000;32:S465–70.
471. Foerster F, Fahrenberg J. Motion pattern and posture: correctly assessed by calibrated accelerometers. *Behav Res Methods Instrum Comput.* 2000;32:450–7.
472. Foerster F, Smeja M, Fahrenberg J. Detection of posture and motion by accelerometry: a validation study in ambulatory monitoring. *Comput Human Behav.* 1999;15:571–83.
473. Nyan M, Tay F, Seah K, Sitoh Y. Classification of gait patterns in the time–frequency domain. *J Biomech.* 2006;39:2647–56.
474. Parkka J, Ermes M, Korpipaa P, Mantyjari J, Peltola J, Korhonen I. Activity classification using realistic data from wearable sensors. *IEEE Trans Inf Technol Biomed.* 2006;10:119–28.
475. Lau H-Y, Tong K-Y, Zhu H. Support vector machine for classification of walking conditions using miniature kinematic sensors. *Med Biol Eng Comput.* 2008;46:563–73.
476. Ng J, Sahakian AV, Swiryn S. Accelerometer-based body-position sensing for ambulatory electrocardiographic monitoring. *Biomed Instrum Technol.* 2003;37:338–46.
477. Rumo M, Amft O, Tröster G, Mäder U. A stepwise validation of a wearable system for estimating energy expenditure in field-based research. *Physiol Meas.* 2011;32:1983–2001.
478. Brage S, Brage N, Franks PW, Ekelund U, Wong M-Y, Andersen LB, Froberg K, Wareham NJ. Branched equation modeling of simultaneous accelerometry and heart rate monitoring improves estimate of directly measured physical activity energy expenditure. *J Appl Physiol.* 2004;96:343–51.
479. Moon JK, Butte NF. Combined heart rate and activity improve estimates of oxygen consumption and carbon dioxide production rates. *J Appl Physiol.* 1996;81:1754–61.
480. Rennie K, Rowsell T, Jebb S, Holburn D, Wareham N. A combined heart rate and movement sensor: proof of concept and preliminary testing study. *Eur J Clin Nutr.* 2000;54:409–14.
481. Edwards AG, Hill JO, Byrnes WC, Browning RC. Accuracy of optimized branched algorithms to assess activity-specific PAEE. *Med Sci Sports Exerc.* 2010;42:672.
482. Haskell WL, Yee MC, Evans A, Irby PJ. Simultaneous measurement of heart rate and body motion to quantitate physical activity. *Med Sci Sports Exerc.* 1993;25:109–15.
483. Luke A, Maki KC, Barkey N, Cooper R, McGEE D. Simultaneous monitoring of heart rate and motion to assess energy expenditure. *Med Sci Sports Exerc.* 1997;29:144–8.
484. Meijer GA, Westerterp KR, Koper H, ten Hoor F. Assessment of energy expenditure by recording heart rate and body acceleration. *Med Sci Sports Exerc.* 1989;21:343–7.



485. Strath SJ, Bassett DR Jr, Thompson DL, Swartz AM. Validity of the simultaneous heart rate-motion sensor technique for measuring energy expenditure. *Med Sci Sports Exerc.* 2002;34:888–94.
486. Strath SJ, Brage S, Ekelund U. Integration of physiological and accelerometer data to improve physical activity assessment. *Med Sci Sports Exerc.* 2005;37: S563–71.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

