

Physical activity characterization: Does one site fit all?

**Clark, C. C. T., Nobre, G. C., Fernandes, J. F. T., Moran, J.,
Drury, B., Mannini, A., Gronek, P. & Podstawski, R.**

Author post-print (accepted) deposited by Coventry University's Repository

Original citation & hyperlink:

Clark, CCT, Nobre, GC, Fernandes, JFT, Moran, J, Drury, B, Mannini, A, Gronek, P & Podstawski, R 2018, 'Physical activity characterization: Does one site fit all?'
Physiological Measurement, vol. 39, no. 9, 09TR02.

<https://dx.doi.org/10.1088/1361-6579/aadad0>

DOI 10.1088/1361-6579/aadad0

ISSN 0967-3334

ESSN 1361-6579

Publisher: IOP Publishing

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

This document is the author's post-print version, incorporating any revisions agreed during the peer-review process. Some differences between the published version and this version may remain and you are advised to consult the published version if you wish to cite from it.

1 **Topical review: Systematic Review article**

2 **Physical Activity Characterisation: Does One Site Fit All?**

3 **Authors**

4 Cain C. T. Clark^{1,2,3*}, Glauber C. Nobre⁴, John F. T. Fernandes³, Jason Moran³, Benjamin
5 Drury³, Andrea Mannini⁵, Piotr Gronek⁶, Robert Podstawski⁷.

6 **Affiliations**

7 ¹ Engineering Behaviour Analytics in Sports and Exercise Research Group, Swansea (U.K.),
8 SA1 8EN.

9 ² School of Life Sciences, Coventry University, Coventry, CV1 5FB, United Kingdom

10 ³ University Centre Hartpury, Higher Education Sport, Gloucestershire (U.K.), GL19 3BE

11 ⁴ Federal Institute of Education, Science and Technology of Ceará, Iguatu, Ceará, Brazil.

12 ⁵ The BioRobotics Institute, Scuola Superiore Sant'Anna, Pisa, Italy.

13 ⁶ Laboratory of Genetics, Department of Dance and Gymnastics, University School of Physical
14 Education in Poznań, Poznań, Poland.

15 ⁷ University of Warmia and Mazury, Olsztyn, Department of Physical Education and Sport
16 Olsztyn, Poland

17 * Corresponding author: Dr. Cain C. T. Clark., Cain.Clark@hartpury.ac.uk.

18 **Key words**

19 Accelerometer; Position; Recognition; Physical Activity; Human Movement; Classification

20 **Abstract**

21 **BACKGROUND**

22 It is evident that a growing number of studies advocate a wrist-worn accelerometer for the
23 assessment of patterns of physical activity *a priori*; yet, the veracity of this site over any other
24 body-mounted location for its accuracy in activity classification is hitherto unexplored.

25 **OBJECTIVE**

26 The objective of this review was to identify the relative accuracy for classifying physical
27 activities according to accelerometer site and analytical technique.

28 **METHODS**

29 A search of electronic databases was conducted using Web of Science, PubMed and Google
30 Scholar. This review included studies written in the English language, published between
31 database inception and December 2017, which characterised physical activities, using a single-
32 accelerometer and reported technique accuracy.

33 **RESULTS**

34 A total of 118 articles were initially retrieved. After duplicates were removed and remaining

1 articles screened, 32 full-text articles were reviewed, resulting in the inclusion of 19 articles
2 that met the eligibility criteria.

3 CONCLUSION

4 There is no ‘one site fits all’ approach to the selection of accelerometer site location or
5 analytical technique. Research design and focus should always inform the most suitable
6 location of attachment, and should be driven by the type of activity being characterised.

7

8 **1 Introduction**

9 In recent years, pervasive, consumer-level wearable physical activity monitors have become
10 commonplace (Thompson, 2016, 2015). With consumerism in mind, manufacturers of such
11 commercial devices offer several wear sites and attachment options, including wrist, hip and
12 pocket attachment. Examples include Fitbit, Polar, Misfit and Jawbone devices (Thompson,
13 2016, 2015). Conversely, until recently, research-grade accelerometers used for empirical
14 investigations, considered to be the *de facto* standard device for objective physical activity
15 monitoring (Mathie *et al.*, 2004; van Hees *et al.*, 2012), were most commonly worn close to
16 the centre of gravity of the body. Devices were traditionally placed at a standardized location,
17 usually at the hip, in order to reflect whole body movement and, thus, energy expenditure
18 (Westerterp, 1999). Recently, however, concerns regarding low compliance to hip-worn
19 accelerometers in empirical studies have resulted, in part, in a shift to utilizing the wrist as the
20 preferred site of attachment (Rowlands *et al.*, 2018). Research-grade accelerometers that are
21 designed to be worn at the wrist are now commonplace (for example; ActiGraph, GENEActiv
22 and Axivity accelerometers). Globally, large-scale, epidemiological projects have advocated
23 the use of accelerometers attached at the wrist, including the National Health and Nutrition
24 Examination Survey (NHANES) in the USA (Freedson and John, 2013), Brazilian birth cohorts
25 (da Silva *et al.*, 2014), the Growing up in Australia Checkpoint (Wake *et al.*, 2014), and
26 Biobank investigations in the UK (<http://www.ukbiobank.ac.uk/about-biobank-uk/>).

27 When utilizing accelerometers for empirical investigations, qualitative evidence suggests that
28 participants prefer a wrist-worn monitor (van Hees *et al.*, 2011; Schaefer *et al.*, 2014), whilst
29 comparable evidence suggests hip-worn monitors are equally desirable (van Hees *et al.*, 2011)
30 and that preference may sometimes depend on the type of activity being undertaken. Offering
31 participants a choice of wear-sites in research studies has been postulated to facilitate greater
32 compliance during a monitoring period, yet this is no more than mere anecdotal reporting.
33 Notwithstanding, although accelerometer output across sites is correlated (Rowlands and

1 Stiles, 2012), the output does differ by wear-site, for example, higher magnitudes are generally
2 found at the wrist vs waist (Rowlands and Stiles, 2012; Hildebrand *et al.*, 2014; Rowlands *et*
3 *al.*, 2014). Hildebrand *et al.* (2014) noted significant differences between the hip and wrist
4 placement in children where acceleration values (physical activity (PA) counts) from the wrist
5 placement, in general, were higher compared with that from the hip. Conversely, Kamada *et*
6 *al.* (2016) highlighted that, in older adults, this difference was insignificant. Similarly,
7 Rowlands *et al.* (2014) concluded that, based on strong linear correlations, output could be
8 predicted from hip- or wrist-worn accelerometer for comparative purposes at the group level.
9 Notwithstanding, it has been asserted that further research needs to be conducted to examine
10 comparisons of specific activities or physical activity intensity levels. Furthermore, due to
11 relatively wide limits of agreement, individual-level comparisons are not, currently,
12 recommended (Mannini *et al.*, 2013; Mannini *et al.*, 2017).

13 Whilst there exists some conjecture as to the optimal wear site for the measurement of PA
14 quantities (PA counts, energy expenditure), this lacks consensus and represents only one,
15 temporal analytical viewpoint. Alternate to traditional temporal analyses; signal processing of
16 accelerometer data has moved beyond the descriptive approach of simply quantifying overall
17 activity using time spent in thresholds or counts per minute (Clark, 2017; Clark *et al.*, 2017b;
18 Clark *et al.*, 2016). It has been shown that more substantive insights are attainable, allowing
19 both quantity and quality to be reported (Chen and Bassett, 2005; Preece *et al.*, 2009; Yang and
20 Hsu, 2010; Clark *et al.*, 2018; Clark *et al.*, 2017a). Recently, Clark *et al.* (2017a) highlighted
21 that the emergence of novel analytical techniques has resulted in the attainment of more
22 sensitive information about physical activity. Further, three separate reviews, (Chen and
23 Bassett (2005), Yang and Hsu (2010) and Clark *et al.* (2017a)) all highlighted that future
24 technological improvements will necessitate examining raw acceleration signals and
25 developing advanced models for accurate energy expenditure prediction and activity
26 classification (Chen and Bassett, 2005; Preece *et al.*, 2009; Yang and Hsu, 2010). In addition
27 to refinement, to account for participant differences, an acceptable level of accuracy for
28 measuring physical activity must be established for analytical techniques, and that ‘qualities’
29 of different activities, such as characteristics of gait, activity duration and idiosyncratic
30 differences be further investigated (Clark *et al.*, 2017a).

1 There have been a number of approaches used for the classification of PA data, starting from a
2 wide set of classification features, then adopting different feature selection/extraction strategies
3 and finally by choosing different automatic classification methodologies (Jain *et al.*, 2000).

4 Pioneering studies in the field attempted the recognition of activities from the output of
5 actigraphy, i.e. activity counts. Such quantities, were obtained by time integrating on fixed time
6 epoch (e.g., 1-min, 1-s) accelerometer recordings with the purpose of estimating the physical
7 activity level of a person (de Vries *et al.*, 2011a). However, the processing of such aggregated
8 quantities only allowed a very rough activity recognition, subsequently, researchers targeted
9 activity recognition by using raw accelerometry data. In fact, raw data retains all of the
10 necessary information for a more accurate and robust activity classification, and has progressed
11 our ability in recognizing even complex vocabularies of activities (Bao and Intille, 2004).
12 Whilst raw acceleration data have been directly processed to recognise activities, *per se*
13 (Hikihara *et al.*, 2014), most of the data reported in the literature were based on the evaluation
14 of accelerometer data for classification features, both in the time and frequency domain. Time
15 domain features are generally obtained by processing windowed portions of the available data.
16 Concerning frequency domain features, fast Fourier transform (FFT) and discrete wavelet
17 transform (DWT) have been used extensively on streaming data. FFT uses the frequency
18 spectrum analysis to distinguish different types of PA (Barralon *et al.*, 2005; Hester *et al.*, 2006;
19 Noury *et al.*, 2004; Mannini *et al.*, 2017; Mannini and Sabatini, 2010; Clark *et al.*, 2017a).
20 Compared with FFT, and its short-time version for time-frequency analysis, DWT allows non-
21 uniform frequency resolution and it has been used in PA studies to detect walking activities
22 based on data collected from hip/lower back accelerometers (Sekine *et al.*, 2002; Barralon *et*
23 *al.*, 2006).

24 An important consideration when classifying data is that large datasets obtained with multiple
25 sensing units will result in multiple features, which necessitates time-consuming data analysis,
26 and may significantly impact the classification methods. In fact, large feature sets may need
27 huge datasets for training computational methods that could be unavailable (the so-called *curse*
28 *of dimensionality*) and, notwithstanding, would slow down the development of the
29 classification system. Consequently, automatic feature selection and feature extraction
30 methods can help reduce the dimensionality of classification problems (Mannini and Sabatini,
31 2010). An example of commonly adopted feature extraction method that remaps features in a
32 different space (that can have reduced dimensionality) is the principal component analysis

1 (PCA), which has been used to define the input of activity classification methods (Long *et al.*,
2 2009).

3 A variety of classification methods have been applied to PA characterisation, ranging from
4 Logistic Regression to Artificial Neural Networks, Support Vector Machines and Decision
5 Trees. Most of the available solutions process data windows independently, whereas in some
6 cases, the time evolution of activities is also used to classify sequences of activities using
7 methods such as Hidden Markov model (HMM) (Mannini and Sabatini, 2010) or Conditional
8 Random Fields (Vinh *et al.*, 2011). Finally, another important aspect to be compared across
9 existing solutions is the cross-validation approach, that is how data are split between subsets
10 for training for machine learning methods and for testing them. The most common approaches
11 are the N-fold cross-validation and the leave-one-subject-out cross validation. Whilst in the
12 former the split is randomized across the full available dataset, in the latter, one participant's
13 data is excluded from training and used for testing. This procedure is repeated to test all subject
14 data, and in doing so, it is possible to simulate the behaviour of the classifier on a new subject's
15 data.

16 Concerning placement site, adjunct to the aforementioned classification methods, there exists
17 no, current, evidence suggesting that one site is better than another. However, previous studies
18 suggest that the wrist is the site that allows higher wear time for long monitoring studies
19 (Freedson and John, 2013; da Silva *et al.*, 2014). Historically speaking, Actigraphy proponents
20 have advocated the hip/pelvis because they are close to the body centre of mass (Westerterp,
21 1999); whilst, distal lower limb sites appear preferred for ambulation-related activities because
22 they are close to the impact site (Mannini and Sabatini, 2010). With further reference to sensor
23 placement site, it is evident that a growing number of studies are advocating a wrist-worn
24 accelerometer *a priori*, yet the veracity of this site over any other body-mounted location for
25 its accuracy in activity recognition and classification is hitherto unexplored. Therefore, the aim
26 of this review is to identify the relative accuracy for classifying physical activities according
27 to accelerometer site, and analytical technique.

28 **2 Methods**

29 We registered this systematic review on the international prospective register of systematic
30 reviews, PROSPERO (<http://www.crd.york.ac.uk/prospero/>): Registration number
31 CRD42018092217. The PRISMA statement for transparent reporting of systematic reviews
32 was followed (Moher *et al.*, 2009).

1 **Eligibility criteria**

2

3 Coding of papers only allowed for studies that attempted to characterise or classify physical
4 activities using one accelerometer and assessed accuracy. We utilized the acronyms PECOT
5 (Population, Exposure, Comparing, Outcome and Type of study) (in accordance with the
6 Preferred Reporting Items for Systematic Reviews and Meta-Analysis protocols; PRISMA-P)
7 to support inclusion of studies. We included studies if they were: (P) Population: children,
8 adolescents, adults and elderly, with positive health conditions; (E) Exposure: accelerometer
9 affixed on the human body; (O) Outcome: relative accuracy for classifying physical activities
10 according to accelerometer site and analytical technique; (T) Type of study: For the purpose of
11 this review, we included observational (e.g., cohort, case-control or cross sectional) and peer
12 reviewed studies.

13 We excluded articles if they: (1) were not human based, or studies that only reported temporal
14 characteristics (e.g. activity counts or energy expenditure); (2) did not report technique
15 accuracy; (3) were technical reports or review articles; (4) case studies, comments, case series,
16 editorial and answers; (5) duplication studies.

17 **Search strategy**

18

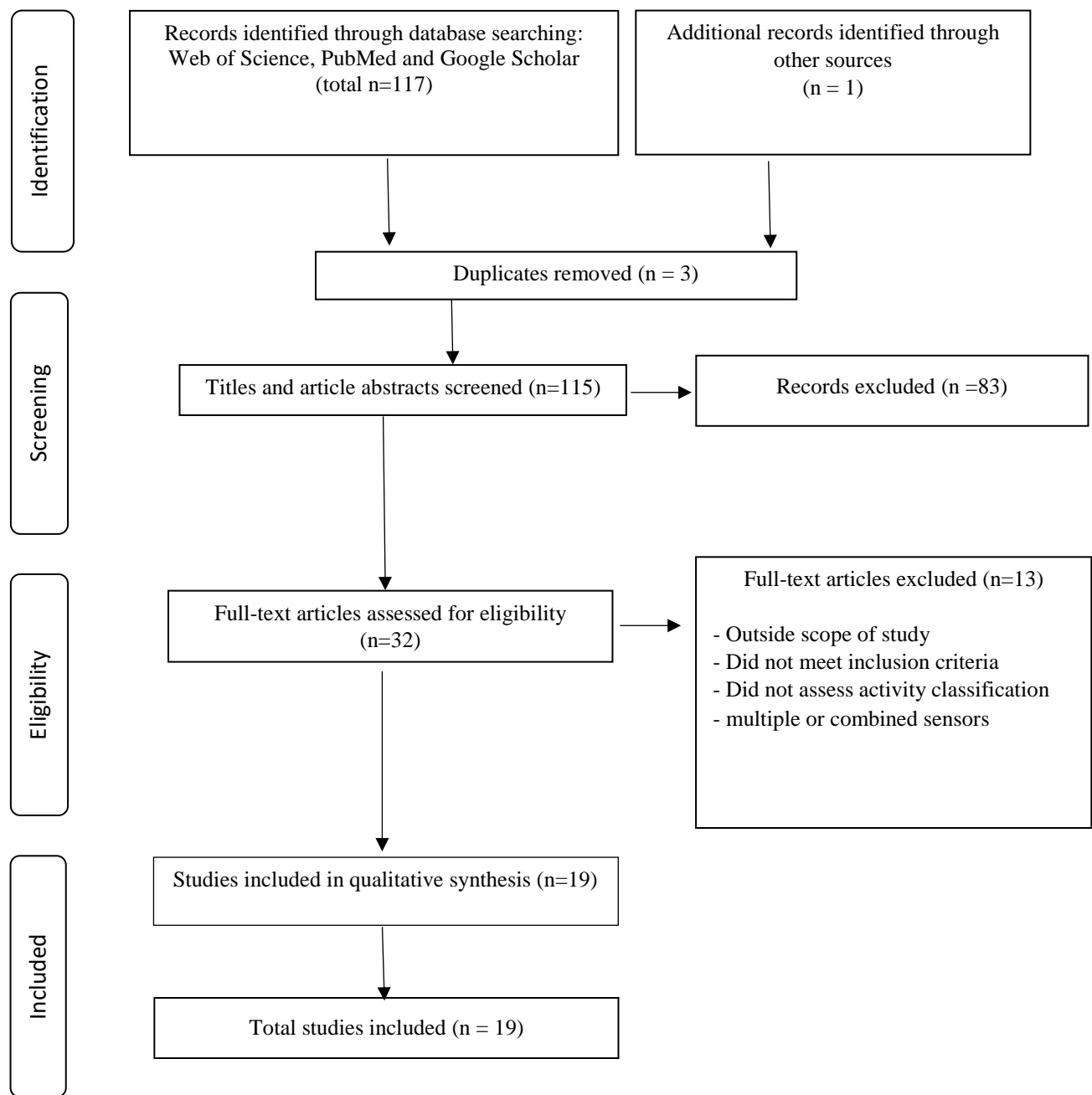
19 In accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis
20 (PRISMA) statement, a computerised search was conducted using the following electronic
21 databases; Web of Science, MEDLINE (accessed by PubMed) and Google Scholar. We used
22 the logic based in specific descriptors, Booleans operators (AND & OR) and help of
23 parentheses. A combination of the following key words were used to locate studies, from
24 database inception to December 2017: ‘accelerometer’, or ‘accelerometry’ or ‘inertial’, and
25 ‘physical activity’, or ‘movement’, or ‘activity’, and ‘classification’, or ‘accuracy’, or
26 ‘identification’. Terms were combined such that every search included one term related to
27 accelerometer, one term related to activity; and one term related to classification. Search string
28 utilised: (Acceleromet* OR inertial) AND ("physical activity" OR movement OR activity)
29 AND (classifi* OR accuracy OR identification OR recogn*). Figure 1 displays the Flow chart
30 of study selection.

31

32

33

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23



24 Figure 1. Flow chart of the search and selection process.

25

26 Data management

27 The exportations of the papers were made in Medline, Ris e Bibtex extension. The data were
28 imported by specific software for systematic reviews, StArt (*State of the Art through Systematic*
29 *Review*) (Fabbri *et al.*, 2016), to facilitate the identification of duplication, excluded and include
30 papers. This procedure was performed by two authors.

1 **Selection process**

2 The selection process sought to identify studies that characterised or classified physical
3 activities using a single-accelerometer unit, assessed the accuracy of the associated analytical
4 technique and were published in the English language between database inception and
5 December 2017. Title and/or abstracts retrieved using the search strategy and those from
6 additional sources were screened independently by two review authors to identify studies that
7 potentially meet the inclusion criteria. The full text of these potentially eligible studies was
8 retrieved and independently assessed for eligibility by two review authors. In instances where
9 the first and second author could not agree, a third, independent, reviewer helped achieve
10 consensus. Duplicates and articles failing to meet inclusion criteria were removed.

11 **Data collection process**

12 Two raters independently extracted data from all articles and reported in the following
13 categories: (A) Aim (activity classified): aim of studies and physical activity measures ; (B)
14 Population: number of subjects, gender, years-old, health conditions and anthropometric
15 characteristics; (C) Device local: human body segment which the accelerometer was worn; (D)
16 Analysis/characterisation technique: computational algorithms utilized in accelerometer signal
17 analysis and classification of physical activity; (E): Overall accuracy: accuracy measure in
18 record movement during exercise; (F) Activity and Device information: type, model,
19 manufacturer, activities classified and recording frequency; (G) Main findings and others
20 relevant information reported by studies. The abstract and full-text of suitable manuscripts were
21 reviewed initially by the same two reviewers, with conclusion supported by a third,
22 independent reviewer.

23 **Outcomes and prioritization**

24 The primary outcome of interest was relative accuracy for classifying physical activities
25 according to accelerometer site. The second outcome are the analytical technique utilized to
26 assessment and classification of physical activities. Given the heterogeneity of samples and
27 instruments observed in studies, we did not load a meta-analysis.

28

29 **Critical appraisal of the included studies**

30 The quality of the studies was appraised using a scale adapted from the ‘Newcastle/Ottawa
31 Scale (NOS) (Wells *et al.*, 2014). Based on the NOS, each study was evaluated using the point
32 system. The authors did critical appraisal of the included studies following these categories:
33 (1) Selection Representativeness of the sample, Sample size, Description of Groups,

1 Ascertainment of exposure; (2) Confound Comparability - Based on design and analysis; (3)
2 Outcome - Assessment of outcome and Statistical test. The assessment considered the follow
3 cut offs: maximum ten score for cross sectional studies. The two raters achieved consensus
4 through discussion. Discrepancies were settled by third author.

5 **3 Results**

6 The electronic search identified 118 potentially relevant articles. Following screening and
7 detailed assessment, 19 studies were deemed suitable for review. (see Figure 1).

8 Table 1 summarises; study aims, population, device locale, analytical technique, overall
9 accuracy, activity and device information and overarching conclusions. Within the 19 studies,
10 included; centre of mass, ankle, and wrist accelerometer positions were investigated. The
11 classification techniques varying across studies (e.g., decision tree, artificial neural network,
12 random forest and hidden Markov model, receive operator characteristic curve analysis,
13 support vector machine (SVM), feed-forward neural networks and k-nearest neighbour).

14 Heterogeneous results were reported in accuracy, varying according device location (between
15 20.35% and 100%). Regarding devices utilised, eight studies (42.1%) utilized the ActiGraph
16 Acc device (de Vries *et al.*, 2011a; De Vries *et al.*, 2011b; Ellis *et al.*, 2014; Ellis *et al.*, 2016;
17 Kuhnhausen *et al.*, 2017; Rowlands *et al.*, 2016; Strath *et al.*, 2015; Trost *et al.*, 2014). Three
18 research groups (15.8%) assessed the physical activity by GENEActiv device (Zhang *et al.*,
19 2012a; Zhang *et al.*; 2012b; Rowlands *et al.*, 2016). Other studies utilized the following
20 accelerometer manufacturers: Moto360 SmartWatch Acc (Amiri *et al.*, 2017), ADXL210E
21 accelerometers (Bao and Intille, 2004), Tracmor Acc (Gyllensten and Bonomi, 2011), Omron
22 Acc (Hikihara *et al.*, 2014), MVN Studi Acc, (Laudanski *et al.*, 2015), Philips NWS Activity
23 Acc (Long *et al.*, 2009), ActiNav system (Mannini and Sabatini, 2011), ST-Microelectronics
24 tri-axial Acc (Oshima *et al.*, 2010) and Minmax tri-axial Acc (Wundersitz *et al.*, 2015).

25 Table 2 (available as supplementary material) shows the critical appraisal of the included
26 studies. The majority (89.5%) reached 70% or superior in Newcastle-Ottawa total score. Two
27 studies scored 100% (Laudanski *et al.*, 2015; Oshima *et al.*, 2010). Only one study accrued a
28 50% NOS total score (Gyllensten and Bonomi, 2011). All studies scored maximum points in
29 outcome criteria (independent assessment and statistical test used to analyse the data was
30 clearly described and appropriate). In contrast, the majority of studies did not highlight
31 representativeness of the sample (63.2%), sample size (63.2%) and description of groups from
32 NOS items for observation research (84.2%).

Table 1. Accelerometer position and classification accuracy

Study	Aim (activity classified)	Population ^a	Device location	Analyses	Overall accuracy	Activity and Device information	Main findings
Amiri et al. (2017)	To detect stereotypic behaviours in children with autism	Fourteen participants, 12 healthy subjects aged between 23-33y, and two subjects (ages 15 and 16) diagnosed with ASD.	Wrist	Features: Time and frequency domain features (FFT, Bispectrum and Wigner Ville Transform). Classifier: Decision tree (best performance) and SVM. Validation: 10-fold cross validation	94.6%	Moto360 SmartWatch Acc, recording at 50 Hz. Stereotypic autistic behaviour was classified.	Wrist worn sensors may be used to monitor stereotypical movements associated with ASD.
Bao and Intille (2004)	To develop and evaluate algorithms to detect physical activities	Twenty participants (no other information provided)	Hip, wrist, arm, ankle and thigh	Features: Time and frequency domain features (FFT). Classifier: SVM. Validation: leave-one-subject- out	Hip: 49.88% Wrist: 32.01% Arm: 20.35% Ankle: 46.92 Thigh: 54.53	ADXL210E accelerometers from Analog Devices, recording at 76.25 Hz. Activities classified; walking, sitting, standing, running, stretching scrubbing, household chores, reading, cycling, strength training, and stair climbing.	Accelerometers can distinguish multiple types of activity. Multi-sensors outperform single units.
de Vries et al. (2011a)	To identify children's physical activity type using ANN models in children.	Fifty-eight children (31 boys and 27 girls, age range = 9-12 y) (age: 11.0 ±0.7y, height: 1.49 ± 0.6 m, weight 42.6 ± 7.8 kg, BMI: 19.0 ±3.0 kg·m ²)	Hip and ankle	Features: statistics extracted form activity counts. Classifier: ANN. Validation: leave-one-subject-out	Uniaxial hip Acc: 72%, triaxial hip Acc: 77%. Uniaxial ankle Acc: 57%, triaxial ankle Acc: 68%.	ActiGraph Acc (uni-and triaxial), recording frequency unspecified, 1-s epochs. Activities classified; sitting, standing, walking, running, rope skipping, playing soccer, and cycling.	ANN models can characterise common physical activities in children. Triaxial Acc are superior to uniaxial Acc.
De Vries et al. (2011b)	To identify types of physical activity in adults using an ANN	Forty-nine subjects (21 men and 28 women) (age 38±11 years,	Hip and ankle	Features: statistics extracted form activity	Hip Acc: 80.4% and ankle Acc:	Actigraph Acc used, recording frequency unspecified, 1-s epochs. Activities classified were; sitting,	ANN models perform well in identifying the type but not the speed of the activity

		height 1.75±0.1 m, weight 73.4±13.3 kg, BMI 23.8±3.4 kg·m ²)		counts. Classifier: ANN. Validation: leave-one-subject-out	77.7% for 5 activities.	standing, using the stairs, and walking and cycling at two self-paced speeds.	of adults from accelerometer data.
Ellis et al. (2014)	To compare accelerometers worn on the wrist and hip for predicting PA type and EE using machine learning	Forty adults (21 women, 19 men). Age: 35.8 ± 12.1 y, BMI: 24.8 ± 2.9 kg·m ²)	Wrist and Hip (analysed individually)	Features: Time and frequency domain features (FFT). Classifier: Random Forest. Validation: leave-one-subject-out	Hip Acc: 92.3%. Wrist Acc: 87.5%, for 4 household activities. Hip Acc: 70.2%. Wrist Acc: 80.2%, for 8 combined household activities.	ActiGraph Acc, recording at 30 Hz, analysed in 1-minute epochs. Eight activities performed and characterised; 1) laundry, 2) window washing, 3) dusting, 4) dishes, 5) sweeping, 6) stairs, 7) walking, 8) running.	Wrist Acc better at predicting activities with significant arm movement. Hip Acc superior for predicting locomotion
Ellis et al. (2016)	To test the performance of machine learning algorithms for classifying PA types from both hip and wrist accelerometer data.	Forty overweight or obese women (age: 55.2 ± 15.3 y; BMI: 32.0 ± 3.7 kg·m ²)	Wrist and hip (analysed individually)	Features: Time and frequency domain features (FFT). Classifier: Random Forest + hidden Markov models for temporal filtering. Validation: leave-one-subject-out	Hip Acc: 89.4%. Wrist Acc: 84.6% for 4 activities.	ActiGraph Acc, recording at 30 Hz. Four activities were; sitting, standing, walking/running, and riding in a vehicle.	Hip and wrist accelerometry can be used to classify free-living activities.
Gyllensten and Bonomi (2011)	To analyse the reproducibility of the accuracy of laboratory-trained classification algorithms in free-living subjects during daily life	20 healthy subjects (10 males) (age: 30 ± 9y; BMI: 23.0 ± 2.6 kg·m ²)	Waist (lower back)	Features: Time and frequency domain features (FFT). Classifier: SVM; feed-forward NN; DT. Validation: leave-one-subject-out	SVM: 95.1%, NN: 91.4%, and DT: 92.2%.	Tracmor Acc, recording at 20 Hz. Activities classified were; lying down, sitting, sitting working at a personal computer, standing still, standing washing dishes, sweeping the floor, walking (treadmill, indoors, outdoors, upstairs, and downstairs), cycling (cycle ergometer and outdoors), and running (treadmill and outdoors).	Capturing daily-life data is essential to training and testing accurate classification models. Combining laboratory and daily-life data to develop classification models should be undertaken.

Hikihara et al. (2014)	To discriminate between non-locomotive and locomotive activities for various physical activities	68 participants 42 boys: 15 (6–9y), and 27 (10–12y). 26 girls: 14 (6–9y), and 12 (10–12y).	Undisclosed	Features: pre-processed acceleration signals. Classifier: threshold-based Validation: no validation	99.1% discriminating between non-locomotive and locomotive activities	Omron Acc, recording at 32 Hz, 10-s epochs. Non-locomotive activities; desk work, Nintendo DS, sweeping up, clearing away, washing the floor, throwing a ball. Locomotive activities; stair descent, stair ascent, normal walking, brisk walking, jogging.	ROC analysis can be used to distinguish between locomotion and non-locomotion.
Kuhnhausen et al. (2017)	To classify children's activities against reference measurements.	70 (43 boys and 27 girls). 8 to 11 years (9.77 ± 0.62 y).	Hip	Features: Time and frequency domain features (FFT).. Classifier: SVM. Validation: leave-one-subject-out	Individual accuracy: 96.9%. Group accuracy: 87.5%.	ActiGraph Acc, recording at 30 Hz. Activities classified; sit, stand, lie, walk, fast run, walk, non-wear.	SVM can provide a reliable method for classifying activities.
Laudanski et al. (2015)	To develop an activity classification algorithm for over-ground walking, stair ascent, and stair descent by individuals presenting with post stroke hemiparesis.	Ten chronic, hemiparetic stroke survivors (67.0 ± 10.9 y)	Shank-mounted	Features: Time and frequency domain features (FFT).. Classifier: SVM. Validation: leave-one-out	100% for three-activities and 94% for five-activities	MVN Studi Acc, recording at 120 Hz. Five gait activities, 1) over-ground walking, 2) stair ascent, and 3) descent, 4) descent with a distinction between stepping pattern used while negotiating stairs (step-over-step(SOS) and 5) step-by-step (SBS)). Only Z-axis component of accelerometer signal analysed. Clinical population. No gender information.	Shank-mounted accelerometry can be used to reliably characterise gait activity in individuals with post stroke hemiparesis
Long et al. (2009)	To characterise daily real-life activities in a naturalistic environment	24 subjects (11 females, ranging in age from 26 to 55 (33.6 ± 7.9 y)).	Hip	Features: Time and frequency domain features (FFT). Spatial domain features defined as sensor orientation variations. PCA is also applied. Classifier: NB,DT. Validation: leave-one-subject-out	NB: 71.5%. DT:72.8%.	Philips NWS Activity Acc, recording at 30 Hz. Activities classified; walking, running, cycling, driving, and sports,	Daily activities can be characterised using different analytical techniques.

Mannini and Sabatini (2011)	To investigate the accuracy of single accelerometer in classifying activity	Six healthy subjects (age: 27.3 ± 2.0 y)	Thigh	and 10-fold cross-validation Features: Time domain features. Classifier: SVM. Validation: individual, leave-one-subject-out and 10-fold cross-validation	Walking: >99%, running >99%.	ActiNav system, recording at 250 Hz. Activities classified; walking and running	A single thigh mounted sensor can be used to accurately distinguish between walking and running.
Oshima et al. (2010)	To develop a new algorithm for classifying physical activity into either locomotive or household activities using a triaxial accelerometer.	Sixty-six volunteers (31 men and 35 women) (age 42.3 ± 13.3 y, height 1.63 ± 0.85 m, weight 61.3 ± 13 kg, BMI 22.7 ± 3.5 kg·m ²)	Hip	Features: Frequency domain features (FFT). Classifier: threshold-based Validation: dataset split in 44 subjects for training and 22 for testing.	63.6%-98.7%	ST-Microelectronics tri-axial Acc, recording at 32 Hz. Twelve physical activities (personal computer work, laundry, dishwashing, moving a small load, vacuuming, slow walking, normal walking, brisk walking, normal walking while carrying a bag, jogging, ascending stairs and descending stairs).	TAU/TAF cut-off value can accurately classify household and locomotive activities.
Rowlands et al. (2016)	To evaluate the accuracy of posture classification using the Sedentary Sphere.	Thirty-four participants (14 male and 20 female patients; age, 27.2 ± 5.9 y; BMI, 23.8 ± 3.7 kg·m ²)	Wrist	Features: statistics extracted from activity counts. Classifier: Threshold based methods Validation: not disclosed	GENEActiv accuracy, sedentary: 74%, postural; 91%. ActiGraph accuracy, sedentary: 75%, postural: 90%.	GENEActiv and ActiGraph Acc, both recording at 100 Hz. Activities classified; sedentary lying, sedentary sitting and upright posture.	These data support the efficacy of the Sedentary Sphere for classification of posture from a wrist-worn accelerometer in adults.
Strath et al. (2015)	To investigate multiple accelerometer sites and their associated PA classification accuracy in adults	Ninety-nine subjects (age: 49 ± 17.4 y, weight: 75.5 ± 16.6 kg)	Ankle, hip and wrist	Features: data representation based on 1s accelerometer epoch. Classifier: SVM	Ankle accuracy: 83%. Hip accuracy: 81.59%. Wrist	ActiGraph Acc, recording frequency unspecified. Activities characterised; six treadmill walking activities, 7 min each. In addition to; daily living, computer work, vacuuming, mopping/sweeping, carrying/lifting boxes of three different weights (4.5, 6.8, and	Whilst accurate in characterising PA, age group specific analyses would be more beneficial.

Trost et al. (2014)	To compare the activity recognition rates of an activity classifier trained on acceleration signal collected on the wrist and hip.	Fifty-two children (28 boys, 24 girls) (age: 13.7±3.1 y, height: 1.60 ±0.15 m, weight: 50.6±13.5 kg)	Hip and wrist	Validation: leave-one-subject-out. Features: Time domain features. Classifier: Logistic regression (LR). Validation: 3-fold modified cross validation (non-overlapping subsets by subject)	accuracy: 69.25% Hip accuracy: 91%, wrist accuracy: 88.4%.	9.1 kg), and walking/intermittent stair climbing, for 7 min each. ActiGraph Acc, recording at 30 Hz. Activities characterised; lying down, hand writing, laundry task, throw and catch, comfortable over-ground walk, and aerobic dance. In addition; computer game, floor sweeping, brisk over-ground walk, basketball, over-ground run/jog, and brisk treadmill walk.	Both the hip and wrist algorithms achieved acceptable classification accuracy, allowing researchers to use either placement for activity recognition.
Wundersitz et al. (2015)	To investigate whether a single wearable tracking device can be used to classify team sport related activities	Seventy-six recreationally active, healthy male participants (age 24.4 ± 3.3 y; height 1.82 ± 0.075 m; mass 77.4 ± 11.6 kg)	Neck, in-line and dorsal to the first to fifth thoracic vertebrae of the spine	Features: Time and frequency domain features (FFT). Features selected by ANOVA and Lasso methods Classifier: RF; SVM; LMT. Validation: hybrid 10-fold/leave-one-out.	The LMT (79-92% classification accuracy) outperformed RF (32-43%) and SVM algorithms (27-40%)	Minmax tri-axial Acc, 100 Hz recording frequency. Each activity circuit included three countermovement jumps, an eight-metre jog, an 8-m COD agility section, two jumps for distance, a 10-m sprint, seven metres of walking, and a tackle bag to be taken to ground with maximum force.	In sporting scenarios where wearable tracking devices are employed, it is both possible and feasible to accurately classify team sport-related activities.
Zhang et al. (2012a)	To classify semi-structured activity	Sixty participants (23 males) (age, 49.4 ± 6.5 y); BMI, 24.6 ± 3.4 kg·m ²)	Right wrist	Features: Time and frequency domain features (FFT + DWT). Classifier: DT, NB,LR,SVM, ANN. Validation: 10-fold cross-validation	94.5%-97.4% (depending on recording frequency)	GENEA Acc used, recording at 5-80Hz. The activities assessed were lying, standing, seated computer work, 4 km.h ⁻¹ walk, 5 km.h ⁻¹ walk, 6 km.h ⁻¹ walk, walking up and down stairs, free-living 6 km/h walk, two household activities (randomly selected from window washing, washing up, shelf stacking, and sweeping), one run (8 km·h ⁻¹ , 10 km·h ⁻¹ or 12 km·h ⁻¹ run), and an optional free-living 10 km·h ⁻¹ run.	Wrist accelerometry for classifying semi-structured activities is accurate, however sampling frequencies >10 Hz and/or more than one axis of measurement are not associated with greater classification accuracy.

Zhang et al. (2012b)	To develop methods to classify physical activities into walking, running, household, or sedentary activities based on raw acceleration data	Sixty participants (23 males) (age: 49.4 ± 6.5 y, BMI: 24.6 ± 3.4 kg·m ²)	Hip and wrist	Features: Time and frequency domain features (FFT + DWT). Classifier: DT, NB,LR,SVM, ANN. Validation: 10-fold cross-validation	Hip accuracy: 99%. Wrist accuracy: 96%.	GENEA Acc used, recording at 80Hz. Activities characterised; sedentary, household, walking, and running.	Wrist accelerometry has good concordance to hip accelerometry for classifying activities.
-----------------------------	---	---	---------------	---	---	--	---

1 Table I definitions; ASD: Autism Spectrum Disorder, ANN: Artificial neural network, ANOVA: Analysis of Variance, Acc: Accelerometer, EE: Energy Expenditure, RF: Random Forest, SVM: Support Vector
2 Machine, LMT: Logistic Model Tree, COD: Change of Direction, SD: Standard Deviation, M: Metres, Y: Years, Kgm²: Kilograms per metres squared, BMI: Body Mass Index, FFT: fast Fourier transformation, PA:
3 Physical Activity, NN: Neural Network, TAU: Total Acceleration Unfiltered, TAF: Total Acceleration Filtered, LR: Logistic Regression, DT: Decision Tree, NB: Bayes Classifier, FFT: fast Fourier transformation,
4 ROC: Receive Operator Characteristic, ^a Age data are mean \pm SD, or range.

5
6
7
8
9
10
11
12
13
14
15
16
17

1 **4 Discussion**

2 The aim of the present systematic review was to identify the relative accuracy for classifying
3 physical activities according to accelerometer site and analytical technique. In accord with the
4 aim of this review, 19 studies that evaluated centre of mass, ankle, and wrist accelerometer
5 positions were reviewed.

6 The modal attachment site of accelerometers for the purpose of characterising physical
7 activities was (circa) the centre of mass, and a range of accuracies and techniques were
8 highlighted (see Table 1). Contemporary studies had a propensity to either focus solely upon,
9 or at least include, wrist-worn accelerometers, and demonstrated comparable accuracy in
10 activity characterisation as hip or centre-of-mass affixed accelerometers. Notwithstanding, the
11 least utilised accelerometer position for physical activity characterisation, in the review, was
12 the lower limbs. Characterising human physical activities and qualities using accelerometers
13 (Umstattd Meyer *et al.*, 2013; Leutheuser *et al.*, 2013; Clark *et al.*, 2018; Clark *et al.*, 2017b;
14 Clark *et al.*, 2017a) demonstrates promising results, with comparably high accuracies between
15 analytical techniques, and across a broad range of activities, including sitting, standing,
16 walking, running, rope skipping, general sports, cycling, general household activities, and stair
17 climbing. In general, studies that compared multiple-sites noted that the hip (centre-of-mass)
18 attachment providing better estimates of whole body activity, wrist attachment provided better
19 estimates in tasks where the centre-of-mass is not mobile, e.g. cycling, whilst ankle attachment
20 facilitated the classification of gait related activity.

21 **4.1 Hip (centre-of-mass) attachment**

22 The most prevalent attachment site for accelerometers, when characterising PA was the aim,
23 was the hip or centre-of-mass. Within this attachment site, ANN, random forest classifier,
24 HMM and SVM were employed. De Vries *et al* (2011a) and De Vries *et al* (2011b) utilised
25 ANN in an attempt to identify sitting, standing, walking, running, soccer and cycling, and
26 general concordance of classification accuracy was reported. De Vries *et al* (2011a)
27 investigated a child population, whilst De Vries *et al* (2011b) utilised middle aged adults,
28 notwithstanding both studies reported accuracies of 77 and 80%, respectively, suggesting that
29 age of the population has little impact on activity classification accuracy.

30 Ellis *et al* (2014) and Ellis *et al* (2016) utilised a random forest classifier, an ensemble machine
31 learning technique, to investigate household and basic postural movements. Ellis *et al* (2014)
32 demonstrated that characterising four basic household movements, an accuracy of 92% was

1 attainable. Similarly, for a comparable number of basic postural movements, Ellis *et al* (2016)
2 reported a characterisation accuracy of ~90% using the random forest classifier technique.
3 However, when more physical activities (eight in total) were added to the classifier matrix,
4 overall accuracy fell by ~20% (Ellis *et al.*, 2014), indicating that there may not be enough
5 degrees of difference between basic movements, when classified from hip mounted
6 accelerometers.

7 Another machine learning algorithm, SVM, has been used extensively with the aim of
8 classifying physical activities in children and adults. Gyllensten and Bonomi (2011) attempted
9 to capture daily-life data in adults, assessing primarily postural-, and some locomotion-based
10 movements. Through the combination of laboratory and daily living measures, respectively, a
11 classification accuracy of over 95% was reported. Similarly, Kuhnhausen *et al* (2017) were
12 able to classify basic postural and locomotor movements, on an individual basis, with an
13 accuracy of over 96%. However, when group-based classifications were attempted, the
14 accuracy fell to 87.5%, highlighting that idiosyncratic differences in even basic movements
15 may hinder sophisticated machine learning techniques. Strath *et al* (2015), whilst investigating
16 an adult population of broad age range (49 ± 17.4 years), showed that a variety of treadmill
17 walking speeds, stair climbing and daily living activities, an overall accuracy of 81.6% was
18 achievable. Impairing the overall accuracy in the study of Strath *et al* (2015), was the large age
19 range of the participants, with the authors asserting that whilst centre of mass mounted
20 accelerometers, combined with SVM, are accurate in characterising PA, age group specific
21 analyses would be more beneficial. Wundersitz *et al* (2015), although not hip mounted, did
22 utilise a central locale in an attempt to classify sport specific activities, including
23 countermovement jumps, walking, jogging, sprinting, change of direction agility drills (COD),
24 jumps for distance, and a tackle bag to be taken to ground with maximum force. It was evident
25 that in sporting scenarios, where wearable tracking devices are employed, whilst it is both
26 possible and feasible to classify team sport-related activities, accuracy, as compared to daily
27 living and basic postural and locomotor movements, is drastically reduced. In contrast to the
28 high level of classification accuracy SVM and random forest classifier techniques display for
29 basic movements, Wundersitz *et al* (2015) reported a range of accuracies between 27-40%, and
30 32-43% for SVM and random forest classifiers, respectively, in recognising a set of 10
31 activities.

32 Feature extraction and recognition of physical activities may take many forms, with
33 probabilistic classifiers (Long *et al.*, 2009), filtered accelerations (Oshima *et al.*, 2010), and

1 FFT (Zhang *et al.*, 2012b) highlighted in this review. Long *et al* (2009) demonstrated that hip
2 mounted accelerometers using Bayes probabilistic classifiers could determine locomotor and
3 sporting activities with accuracies up to 71.5% in adults. Whilst in a comparable population,
4 Oshima *et al* (2010) reported accuracies ranging from 63-98%, concluding that the more
5 complex movement, the less accurate feature extraction may be. In a child population, Trost *et*
6 *al* (2014) found that hip (accelerometer) algorithms achieved good classification accuracy
7 (91%) for basic postural and locomotor activities. Finally, Zhang *et al* (2012b) extracted
8 features of activity through the use of FFT, with focus on basic distinction between
9 sedentarism, household activity, walking and running, and reported classification accuracy of
10 99%.

11
12 It was evident that daily activities can be characterised using different analytical techniques,
13 with a hip or centrally placed accelerometer. A common theme that has emerged from this
14 review is that basic postural and locomotor type distinction is very accurate, with relative
15 accuracies of 77-99%. However, when more complex, or indeed intense, movements are
16 investigated, accuracy deleteriously suffers, with accuracies dropping as low as 27%. Many of
17 these studies failed to account, or indeed, explicitly report, anthropometric and physiological
18 metrics such as age, sex and fitness, which could conceivably affect patterns of movement.
19 Concerningly, a confounding limitation was that key accelerometer information was omitted
20 in many cases, such as recording frequencies utilised and band-pass filters applied, which are
21 vital pieces of information to standardise and replicate analyses. Although substantial gains
22 have been made utilising classification analytics to develop deeper insights into human
23 physical activity data, the underlying algorithms require further development. Further, work
24 surrounding age and population specific algorithms need to be addressed, it is evident that
25 analyses that were able to take into account idiosyncratic movement patterns were consistently
26 accurate, yet in many cases, when analyses were scaled up to group or cohort-based analyses,
27 accuracy is reduced.

28 **4.2 Leg attachment**

29 Similar to centrally located accelerometer sites, analytical techniques employed for
30 accelerometers attached to the leg included; ANN, FFT, and SVM. In children, De Vries *et al*
31 (2011a) highlighted that ankle-affixed accelerometers were less able to accurately classify
32 sitting, standing, walking, running, rope skipping, playing soccer, and cycling, collectively,
33 than centrally located monitors, 68 vs. 77%, respectively. Interestingly, De Vries *et al* (2011a)

1 reported better accuracy in identifying locomotion/gait with a hip vs ankle accelerometer; this
2 is in contrast to previous reports. Barnes *et al.* (2017) systematically demonstrated the
3 capability of an ankle mounted accelerometer to distinguish specific portions of gait and
4 ambulation, whilst Mannini *et al.* (2013) highlighted that for movement quality characteristics
5 related to ambulation, an ankle-mounted monitor is most suitable. This discrepancy was
6 highlighted by the authors (De Vries *et al.*, 2011a), and attributed to the low sampling fidelity,
7 where data were transformed into 1-s epochs, opposed to maintaining the raw data. In adults,
8 De Vries *et al.* (2011b) noted ~78% accuracy in identifying postural movement and cycling,
9 which was less accurate than a centrally placed monitor, identifying the same activities (80%).
10 It was evident (De Vries *et al.*, 2011a; De Vries *et al.*, 2011b) that using ANN and leg-mounted
11 accelerometers for identifying the type of activity demonstrates good accuracy, however, the
12 speed of the activity of adults from accelerometer data is less distinguishable. Notwithstanding,
13 a common limitation of De Vries *et al.* (2011a) and De Vries *et al.* (2011b) is that the raw
14 acceleration data (both 30 data points per second) were filtered into 1-s epochs, thereby
15 reducing the sensitivity and ability of the analyses to differentiate specific portions of gait or
16 locomotion.

17 Whilst most studies included in this review focussed on healthy individuals, Laudanski *et al.*
18 (2015) utilised hemiparetic stroke surviving participants and assessed five specific locomotor
19 activities. In contrast to De Vries *et al.* (2011a) and De Vries *et al.* (2011b), who reduced raw
20 acceleration data, Laudanski *et al.* (2015) recorded at a high frequency (120 Hz) and did not
21 artificially reduce the data points. The resultant analyses, FFT, yielded a classification accuracy
22 of 100% and 94% for three and five activities, respectively. Notwithstanding, however,
23 limitations in Laudanski *et al.* (2015) included a paucity of sex information and no exact
24 description of the accelerometer placement. The final study utilising an ankle-mounted
25 accelerometer to classify PA, in this review, is Strath *et al.* (2015), who utilised SVM and
26 showed that in a variety of treadmill walking speeds, stair climbing and daily living activities,
27 an overall accuracy of 83% was achievable. Whilst Strath *et al.* (2015) included no recording
28 frequency or data reduction information, given that the ankle- outperformed the hip
29 accelerometer in locomotor activity classification (83 vs 81%), this suggests that the raw
30 acceleration data were utilised and not reduced. Strath *et al.* (2015) asserted that whilst accurate
31 in characterising PA, age group specific analyses are necessitated. Furthermore, whilst age
32 specific analyses could be used, maturation specific may be more relevant and accurate i.e. a
33 different approach must be taken in youth than in adults.

1 It was apparent that for whole body movements, leg mounted accelerometers were
2 systematically less accurate than hip or centrally located devices, and given the well-
3 established link between centrally located accelerometers and energy expenditure estimation
4 (Siervo *et al.*, 2013), it is unsurprising that monitors attached to the leg were less accurate.
5 Conversely, however, when locomotion, or gait, was the focus of analyses, leg mounted
6 monitors were more consistently accurate, highlighting that one (accelerometer) site, does not
7 fit all. In some cases, however, excessive data reduction (i.e. reducing 30 data points per second
8 to one) led to hip- outperforming leg-mounted accelerometers for locomotor classification
9 accuracy. Furthermore, limitations present across studies included limited reporting of sex
10 information, accelerometer recording frequencies and band-pass filtering applied, thereby
11 confounding researchers' abilities to replicate and discuss data analyses.

12 **4.3 Wrist attachment**

13 In recent years, the tendency for empirical investigations to incorporate wrist-worn
14 accelerometry has grown, and within such investigations, analytical techniques have remained
15 consistent between central, leg and wrist accelerometer sites. In this review, studies utilising
16 wrist mounted accelerometers used SVM, FFT, random forest classifier and HMM.

17 Amiri *et al* (2017) sought to detect stereotypic behaviours in children with autism, including
18 flapping, sibbing (self-hitting) and painting with an accuracy of 94%. Given the propensity of
19 ASD sufferers to gesticulate with their arms, a wrist-mounted accelerometer was asserted to be
20 a logical placement (Amiri *et al.*, 2017), giving further credence to a one site *does not* fit all
21 approach. Strath *et al* (2015) utilised SVM and showed that in a variety of treadmill walking
22 speeds, stair climbing and daily living activities, an overall accuracy of ~70% was achievable.
23 This was, however, markedly lower than both hip and ankle monitors used to characterise the
24 same movements (~70 vs 81 and 83%, respectively). Trost *et al* (2014) compared wrist and hip
25 accelerometer placements, combined with feature extraction methods, to classify sedentary and
26 daily life behaviours, and asserted that both the hip and wrist algorithms provide acceptable
27 classification accuracy, allowing researchers to use either placement for activity recognition.
28 The relative accuracies of hip and wrist monitors were 91% and 88%, respectively,
29 demonstrating very good agreement between the two placements. The concordance in
30 accuracies reported by Trost *et al* (2014) are contrary to Strath *et al* (2015) and this is likely
31 due to the former focussing on whole-body type activity whilst the latter predominantly
32 analysed locomotion types.

1 Ellis *et al* (2014) and Ellis *et al* (2016) reported that wrist-mounted accelerometers were less
2 accurate than hip-mounted counterparts when characterising four, mainly locomotor
3 movements (household, stairs, walking, running), in children and adults, were analysed. Ellis
4 *et al* (2014) and Ellis *et al* (2016) demonstrated accuracies of 87 and 84, respectively for wrist
5 monitors, vs. 92 and 89%, respectively for hip monitors. However, in Ellis *et al* (2014), when
6 more activities (four wrist-dominant activities; laundry, window washing, dusting, dish
7 cleaning) were added into the random forest classifier, this led to the wrist- outperforming the
8 hip-mounted monitor (80% vs. 70%, respectively). Zhang *et al* (2012a) investigated postural
9 and locomotor data from single, dual, and three axes at sampling rates of 5, 10, 20, 40, and 80
10 Hz, where mathematical models based on features extracted from mean, standard deviations,
11 FFT, and wavelet decomposition were built. Zhang *et al* (2012a) reported high classification
12 accuracy, irrespective of the number of accelerometer axes for data collected at 80 Hz (97%),
13 40 Hz (97%), 20 Hz (97%), and 10 Hz (97%) and 5 Hz (95%). The authors further asserted
14 that sampling frequencies of 10 Hz and/or more than one axis of measurement were not
15 associated with greater classification accuracy. Utilising the same analytical approach as Zhang
16 *et al* (2012a) (i.e. FFT and wavelet decomposition), Zhang *et al* (2012b) showed that when
17 classifying sedentary, household, walking, and running behaviours, wrist accelerometry has
18 good concordance to hip accelerometry, 96 vs 99%, respectively. Notwithstanding, it was
19 concluded in both, Zhang *et al* (2012a) and Zhang *et al* (2012b), that classification of a small
20 number of basic postural and locomotor movements is reliable and accurate, yet further work
21 and refinement, considering a wider range of activities, is necessary. The final study included
22 in this review, Rowlands *et al* (2016), pioneered a novel analytical approach to activity
23 classification; a threshold-based concept coined the ‘sedentary sphere’. The algorithm and cut-
24 points developed were primarily targeted at classifying postural movement and changes, and
25 as such, noted an accuracy of ~90%, irrespective of accelerometer type.

26 One site is clearly not suitable for all types of PA assessment, whether the activity being
27 classified is locomotor (or gait) based, whole body, or wrist dominant has large impact on the
28 capability of single-position accelerometers to accurately characterise activity, regardless of
29 analytical technique employed. When activities go beyond basic postural or locomotor
30 movements, wrist-mounted accelerometers, irrespective of analytical technique, lose efficacy
31 in such tasks.

32 **4.4 Considerations**

1 Although comparison of overall accuracy of physical activity characterisation provides an
2 overarching view of the literature, precautions must be taken when interpreting this, or indeed
3 any, accuracy-based metric. For instance, the dataset, i.e., the number of activity classes and
4 amount of data classified must be considered. If the dataset is complex involving many classes,
5 for example in De Vries *et al* (2011a) and De Vries *et al* (2011b), whom attempted to identify
6 sitting, standing, walking, running, soccer and cycling, a 10% gross-error is contrasting from
7 a 10% error in a two-class problem, as in Mannini and Sabatini (2011) whom distinguished
8 between walking and running. Moreover, if an investigation is designed on a complex dataset,
9 the accuracy meaning is necessarily different from that inferred from comparably simplistic
10 studies. Likewise, the activity type is also important; empirical investigations that target very
11 simple, usually postural, activities, like sit-stand-walk, will necessarily be more likely to
12 characterise, accurately, the activity, whereas those studies targeting very complex sets, such
13 as Bao and Intille (2004), conversely, require a far more robust and comprehensive set of
14 computational methods.

15 A further consideration, and indeed avenue for further research, is the classifier training and
16 validation approach utilised, for example; 60%-40% train test split, n-fold cross-validation,
17 individual validation, leave-one-subject-out cross-validation. In fact, by using a validation
18 approach that uses data from a specific subject that is used also for testing, greater accuracy is
19 more easily attainable than using leave-one-subject out cross validation which means testing
20 the solution on a subject that was excluded from training. Further work must also be conducted,
21 examining the influence of time domain, frequency domain, PCA-based statistical analyses.
22 The varying analytics and outputs relate to computational complexity of the solutions; complex
23 features can translate into complex data processing that could, therefore, limit the real-time
24 capabilities of the method.

25 The accelerometer brand-specific details and operational capabilities and/or set-up need to be
26 concomitantly considered. Evident in this review, band-pass filtering applied to raw
27 accelerometer is, for the most part, unreported; accelerometer-recording frequencies (ranging
28 from 10-120 Hz) are varied, inconsistent and, often, not rationalised against the expected
29 frequency of human activity. The Nyquist-Shannon sampling theorem specifies that the sample
30 must contain all the available frequency information from the signal to result in a faithful
31 reproduction of the analogue waveform signal. Further, put simply, if the highest frequency
32 component, in Hz, for a given analogue signal is f_{\max} , according to the Nyquist-Shannon
33 sampling Theorem, the sampling rate must be at least $2f_{\max}$, or twice the highest analogue

1 frequency component. If the sampling rate is less than $2f_{\max}$, and/or if all the available frequency
2 information is not available, the signal will not be correctly represented in the digitized output
3 (Shannon, 1949; Farrow *et al.*, 2011).

4 Practically, most of the state of the art solutions adopt machine learning algorithms for which
5 the computational cost is mainly associated with the training phase. The test phase most
6 commonly compatible with online solutions that process data as soon as they are available by
7 the wearable sensor. The only limitations are related to the fact that acceleration signals are
8 processed on a window-by-window base and then, a delay related to the filling of the data
9 window and to the features calculation is usually present, notwithstanding, this should not
10 generally prevent any researcher from implementing online working solutions.

11 Overall, these considerations must be appreciated when researchers, practitioners and
12 clinicians are attempting to characterise physical activities. Furthermore, the expert-led
13 consensus of van Hees *et al.* (2016) asserts authors must; consider and present sensor
14 specification, algorithms for data processing must be, comprehensively, reported, and
15 important decisions regarding empirical aims, motivations and expected outcomes must be
16 established and documented, thereby facilitating interpretation.

17 **5 Conclusion**

18 The aim of the present review was to identify the relative accuracies for classifying physical
19 activities according to accelerometer site and analytical technique. In accord with the aim of
20 this review, it was found that overall accuracy for various accelerometer sites for activity
21 recognition were comparable; accelerometer site for activity recognition needs to be carefully
22 considered based on the type of activity being investigated; and that there is no ‘one site fits
23 all’ approach to accelerometer site, or indeed analytical technique,. Whilst contemporary
24 research has cited accelerometer position as a rationale for poor compliance, this is
25 inconclusive and not a logical reason in itself to favour one position *a priori*, further
26 confounding, the current literature pool is bereft of uniformity and consistency. We therefore
27 believe that this systematic review will provide practical information and guidance to current
28 and prospective researchers. Research design and focus should always inform the most suitable
29 location of attachment, and driven by what type of activity is being characterised. Furthermore,
30 in line with the expert consensus of van Hees *et al* (2016), detailed specification of sensors
31 need to be routinely provided and each fundamental step of algorithms for processing raw

1 accelerometer data need to be documented and motivated, to facilitate interpretation,
2 replication and discussion.

3 **Compliance with Ethical Standards**

4 **Funding**

5 No sources of funding were used to assist in the preparation of this article

6 **Conflict of Interests**

7 The authors do hereby declare that they have no conflicts of interest relevant to the content of
8 this review.

9 **References**

- 10 Amiri A M, Peltier N, Goldberg C, Sun Y, Nathan A, Hiremath S V and Mankodiya K 2017
11 WearSense: Detecting Autism Stereotypic Behaviors through Smartwatches
12 *Healthcare (Basel)* **5**
- 13 Bao L and Intille S S 2004 Activity Recognition from User-Annotated Acceleration Data.
14 (Berlin, Heidelberg: Springer Berlin Heidelberg) pp 1-17
- 15 Barnes C M, Clark C C, Holton M D, Stratton G and Summers H D 2017 Quantitative Time
16 Profiling of Children's Activity and Motion *Medicine and science in sports and exercise*
17 **49** 183-90
- 18 Barralon O, Vuillerme N and Noury N 2006 Walk detection with a kinematic sensor: frequency
19 and wavelet comparison. In: *Proceedings of the 28th IEEE EMBS Annual International*
20 *Conference*, (New York, NY: Engineering in Medicine and Biology Society) pp 1711-
21 4
- 22 Barralon P, Noury N and Vuillerme N 2005 Classification of daily physical activities from a
23 single kinematic sensor *Conference proceedings : ... Annual International Conference*
24 *of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in*
25 *Medicine and Biology Society. Conference* **3** 2447-50
- 26 Chen K Y and Bassett D R, Jr. 2005 The technology of accelerometry-based activity monitors:
27 current and future *Medicine and science in sports and exercise* **37** S490-500
- 28 Clark C C, Barnes C M, Stratton G, McNarry M A, Mackintosh K A and Summers H D 2017a
29 A Review of Emerging Analytical Techniques for Objective Physical Activity
30 Measurement in Humans *Sports medicine* **47** 439-47
- 31 Clark C C T 2017 Profiling movement and gait quality using accelerometry in children's
32 physical activity: consider quality, not just quantity *British journal of sports medicine*
- 33 Clark C C T, Barnes C M, Holton M D, Mackintosh K A, Summers H D and Stratton G 2016
34 Profiling movement quality characteristics according to body-mass index in children.
35 In: *21st Annual Congress of the European College of Sport Science*, (Vienna, Austria
- 36 Clark C C T, Barnes C M, Summers H D, Mackintosh K A and Stratton G 2018 Profiling
37 movement quality characteristics of children (9-11y) during recess *European Journal*
38 *of Human Movement* **39** 143-60
- 39 Clark C C T, Barnes C M, Swindell N J, Holton M D, Bingham D D, Collings P J, Barber S E,
40 Summers H D, Mackintosh K A and Stratton G 2017b Profiling Movement and Gait
41 Quality Characteristics in Pre-School Children *Journal of motor behavior* 1-9

- 1 da Silva I C, van Hees V T, Ramires V V, Knuth A G, Bielemann R M, Ekelund U, Brage S
2 and Hallal P C 2014 Physical activity levels in three Brazilian birth cohorts as assessed
3 with raw triaxial wrist accelerometry *International journal of epidemiology* **43** 1959-
4 68
- 5 de Vries S I, Engels M and Garre F G 2011a Identification of children's activity type with
6 accelerometer-based neural networks *Medicine and science in sports and exercise* **43**
7 1994-9
- 8 De Vries S I, Garre F G, Engbers L H, Hildebrandt V H and Van Buuren S 2011b Evaluation
9 of neural networks to identify types of activity using accelerometers *Medicine and*
10 *science in sports and exercise* **43** 101-7
- 11 Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D and Marshall S 2014 A random forest
12 classifier for the prediction of energy expenditure and type of physical activity from
13 wrist and hip accelerometers *Physiological measurement* **35** 2191-203
- 14 Ellis K, Kerr J, Godbole S, Staudenmayer J and Lanckriet G 2016 Hip and Wrist Accelerometer
15 Algorithms for Free-Living Behavior Classification *Medicine and science in sports and*
16 *exercise* **48** 933-40
- 17 Fabbri S, Octaviano F, Silva C, Di Thommazo A, Hernandez E and Belgamo A 2016
18 Improvements in the Start tool to better support the systematic review process. In: *20th*
19 *International Conference on Evaluation and Assessment in Software Engineering*
20 (Limerick, Ireland)
- 21 Farrow C L, Shaw M, Kim H, Juhás P and Billinge S J 2011 Nyquist-Shannon sampling
22 theorem applied to refinements of the atomic pair distribution function *Physical Review*
23 *B* **84** 134105
- 24 Freedson P S and John D 2013 Comment on "estimating activity and sedentary behavior from
25 an accelerometer on the hip and wrist" *Medicine and science in sports and exercise* **45**
26 962-3
- 27 Gyllensten I C and Bonomi A G 2011 Identifying types of physical activity with a single
28 accelerometer: evaluating laboratory-trained algorithms in daily life *IEEE transactions*
29 *on bio-medical engineering* **58** 2656-63
- 30 Hester T, Sherril D M, Hamel M, Perreault K, Boissy P and Bonato P 2006 Identification of
31 tasks performed by stroke patients using a mobility assistive device *Conference*
32 *proceedings : ... Annual International Conference of the IEEE Engineering in Medicine*
33 *and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*
34 **1** 1501-4
- 35 Hikiyara Y, Tanaka C, Oshima Y, Ohkawara K, Ishikawa-Takata K and Tanaka S 2014
36 Prediction models discriminating between nonlocomotive and locomotive activities in
37 children using a triaxial accelerometer with a gravity-removal physical activity
38 classification algorithm *PloS one* **9** e94940
- 39 Hildebrand M, VT V A N H, Hansen B H and Ekelund U 2014 Age group comparability of
40 raw accelerometer output from wrist- and hip-worn monitors *Medicine and science in*
41 *sports and exercise* **46** 1816-24
- 42 Jain A K, Duin R and Mao J 2000 Statistical pattern recognition: A review *IEEE Transactions*
43 *on Automatic Pattern Analysis and Machine Intelligence* **22** 4-37
- 44 Kamada M, Shiroma E J, Harris T B and Lee I M 2016 Comparison of physical activity
45 assessed using hip- and wrist-worn accelerometers *Gait & posture* **44** 23-8
- 46 Kuhnhausen J, Dirk J and Schmiedek F 2017 Individual classification of elementary school
47 children's physical activity: A time-efficient, group-based approach to reference
48 measurements *Behavior research methods* **49** 685-97
- 49 Laudanski A, Brouwer B and Li Q 2015 Activity classification in persons with stroke based on
50 frequency features *Medical engineering & physics* **37** 180-6

- 1 Leutheuser H, Schuldhaus D and Eskofier B M 2013 Hierarchical, Multi-Sensor Based
2 Classification of Daily Life Activities: Comparison with State-of-the-Art Algorithms
3 Using a Benchmark Dataset *PloS one* **8** e75196
- 4 Long X, Yin B and Aarts R M 2009 Single-accelerometer-based daily physical activity
5 classification *Conference proceedings : ... Annual International Conference of the*
6 *IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and*
7 *Biology Society. Conference* **2009** 6107-10
- 8 Mannini A, Intille S S, Rosenberger M, Sabatini A M and Haskell W 2013 Activity recognition
9 using a single accelerometer placed at the wrist or ankle *Medicine and science in sports*
10 *and exercise* **45** 2193-203
- 11 Mannini A, Rosenberger M, Haskell W L, Sabatini A M and Intille S S 2017 Activity
12 Recognition in Youth Using Single Accelerometer Placed at Wrist or Ankle *Medicine*
13 *and science in sports and exercise* **49** 801-12
- 14 Mannini A and Sabatini A M 2010 Machine learning methods for classifying human physical
15 activity from on-body accelerometers *Sensors* **10** 1154-75
- 16 Mannini A and Sabatini A M 2011 On-line classification of human activity and estimation of
17 walk-run speed from acceleration data using support vector machines *Conference*
18 *proceedings : ... Annual International Conference of the IEEE Engineering in Medicine*
19 *and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*
20 **2011** 3302-5
- 21 Mathie M J, Coster A C, Lovell N H and Celler B G 2004 Accelerometry: providing an
22 integrated, practical method for long-term, ambulatory monitoring of human movement
23 *Physiological measurement* **25** R1-20
- 24 Moher D, Liberati A, Tetzlaff J, Altman D G and Group P 2009 Preferred reporting items for
25 systematic reviews and meta-analyses: the PRISMA statement *PLoS medicine* **6**
26 e1000097
- 27 Noury N, Barralon P, Couturier P, Favre-Reguillion F, Guillemaud R, Mestais C, Caritu Y,
28 David D, Moine S, Franco A, Guiraud-By F, Berenguer M and Provost H 2004 A
29 microsystem based on MEMS for activity monitoring of the frail elderly in their daily
30 life *Conference proceedings : ... Annual International Conference of the IEEE*
31 *Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and*
32 *Biology Society. Conference* **5** 3305 - 8
- 33 Oshima Y, Kawaguchi K, Tanaka S, Ohkawara K, Hikiyama Y, Ishikawa-Takata K and Tabata
34 I 2010 Classifying household and locomotive activities using a triaxial accelerometer
35 *Gait & posture* **31** 370-4
- 36 Preece S J, Goulermas J Y, Kenney L P and Howard D 2009 A comparison of feature extraction
37 methods for the classification of dynamic activities from accelerometer data *IEEE*
38 *transactions on bio-medical engineering* **56** 871-9
- 39 Rowlands A V, Harrington D M, Bodicoat D H, Davies M J, Sherar L B, Gorely T, Khunti K
40 and Edwardson C L 2018 Compliance of Adolescent Girls to Repeated Deployments
41 of Wrist-worn Accelerometers *Medicine and science in sports and exercise*
- 42 Rowlands A V, Rennie K, Kozarski R, Stanley R M, Eston R G, Parfitt G C and Olds T S 2014
43 Children's physical activity assessed with wrist- and hip-worn accelerometers *Medicine*
44 *and science in sports and exercise* **46** 2308-16
- 45 Rowlands A V and Stiles V H 2012 Accelerometer counts and raw acceleration output in
46 relation to mechanical loading *Journal of biomechanics* **45** 448-54
- 47 Rowlands A V, Yates T, Olds T S, Davies M, Khunti K and Edwardson C L 2016 Sedentary
48 Sphere: Wrist-Worn Accelerometer-Brand Independent Posture Classification
49 *Medicine and science in sports and exercise* **48** 748-54

- 1 Schaefer S E, Van Loan M and German J B 2014 A feasibility study of wearable activity
2 monitors for pre-adolescent school-age children *Preventing chronic disease* **11** E85
- 3 Sekine M, Tamura T, Akay M, Fujimoto T, Togawa T and Fukui Y 2002 Discrimination of
4 walking pattern using wavelet-based fractal analysis *IEEE transactions on neural
5 systems and rehabilitation engineering : a publication of the IEEE Engineering in
6 Medicine and Biology Society* **10** 188-96
- 7 Shannon C 1949 Communication in the presence of noise *Proceedings of the Institute of Radio
8 Engineers* **37** 10-21
- 9 Siervo M, Bertoli S, Battezzati A, Wells J C, Lara J, Ferraris C and Tagliabue A 2013 Accuracy
10 of predictive equations for the measurement of resting energy expenditure in older
11 subjects *Clinical nutrition*
- 12 Strath S J, Kate R J, Keenan K G, Welch W A and Swartz A M 2015 Ngram time series model
13 to predict activity type and energy cost from wrist, hip and ankle accelerometers:
14 implications of age *Physiological measurement* **36** 2335-51
- 15 Thompson W R 2015 Worldwide survey of fitness trends for 2016 *ACSM's health & fitness
16 journal* **19** 9-18
- 17 Thompson W R 2016 Worldwide survey of fitness trends for 2017 *ACSM's health & fitness
18 journal* **20** 8-17
- 19 Trost S G, Zheng Y and Wong W K 2014 Machine learning for activity recognition: hip versus
20 wrist data *Physiological measurement* **35** 2183-9
- 21 Umstatted Meyer M R, Baller S L, Mitchell S M and Trost S G 2013 Comparison of 3
22 accelerometer data reduction approaches, step counts, and 2 self-report measures for
23 estimating physical activity in free-living adults *Journal of physical activity & health*
24 **10** 1068-74
- 25 van Hees V T, Gorzelniak L, Leon E, Eder M, Pias M, Taherian S, Ekelund U, Renstrom F,
26 Franks P, Horsch A and Brage S 2012 A method to compare new and traditional
27 accelerometry data in physical activity monitoring. In: *World of Wireless Mobile and
28 Multimedia Networks (WoWMoM)*, (Montreal, QC, Canada pp 1-6
- 29 van Hees V T, Renstrom F, Wright A, Gradmark A, Catt M, Chen K Y, Lof M, Bluck L,
30 Pomeroy J, Wareham N J, Ekelund U, Brage S and Franks P W 2011 Estimation of
31 daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-
32 axial accelerometer *PloS one* **6** e22922
- 33 van Hees V T, Thaler-Kall K, Wolf K H, Brond J C, Bonomi A, Schulze M, Vigl M, Morseth
34 B, Hopstock L A, Gorzelniak L, Schulz H, Brage S and Horsch A 2016 Challenges and
35 Opportunities for Harmonizing Research Methodology: Raw Accelerometry *Methods
36 of information in medicine* **55** 525-32
- 37 Vinh L T, Lee S, Le H X, Ngo H Q, Kim H I, Han M and Lee Y-K 2011 Semi-Markov
38 conditional random fields for accelerometer-based activity recognition *Applied
39 Intelligence* **35** 226-41
- 40 Wake M, Clifford S, York E, Mensah F, Gold L H, Burgner D, Davies S, Azzopardi P, Baur
41 L, Carlin J, Cheung M, Crist G, Dwyer T, Edwards B, Olds T, Ranganathan S, Rogers
42 H, Simm P, Saffrey R, Sawyer R G, Stevens L, Wong T and Zubrick S 2014 Introducing
43 growing up in Australia's child health checkpoint: A physical health and biomarkers
44 module for the longitudinal study of Australian children *Fam Matters* 15-23
- 45 Wells G A, Shea B, O'Donnell D, Peterson J, Welch V, Losos M and Tugwell P 2014 The
46 Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in
47 meta-analyses.
- 48 Westerterp K R 1999 Physical activity assessment with accelerometers *International journal
49 of obesity and related metabolic disorders : journal of the International Association for
50 the Study of Obesity* **23 Suppl 3** S45-9

1 Wundersitz D W, Josman C, Gupta R, Netto K J, Gastin P B and Robertson S 2015
2 Classification of team sport activities using a single wearable tracking device *Journal*
3 *of biomechanics* **48** 3975-81
4 Yang C C and Hsu Y L 2010 A review of accelerometry-based wearable motion detectors for
5 physical activity monitoring *Sensors* **10** 7772-88
6 Zhang S, Murray P, Zillmer R, Eston R G, Catt M and Rowlands A V 2012a Activity
7 classification using the GENE: optimum sampling frequency and number of axes
8 *Medicine and science in sports and exercise* **44** 2228-34
9 Zhang S, Rowlands A V, Murray P and Hurst T L 2012b Physical activity classification using
10 the GENE wrist-worn accelerometer *Medicine and science in sports and exercise* **44**
11 742-8

12

13

14

15