Gait & Posture 61 (2018) 98–110

Contents lists available at ScienceDirect

Gait & Posture

journal homepage: www.elsevier.com/locate/gaitpost

Review

Calibration of raw accelerometer data to measure physical activity: A systematic review

Márcio de Almeida Mendes[a,*], Inácio C.M. da Silva[a], Virgílio V. Ramires[a], Felipe F. Reichert[b], Rafaela C. Martins[a], Elaine Tomasi[a,*]

[a] Post-Graduate Program in Epidemiology, Federal University of Pelotas, Brazil
[b] School of Physical Education, Federal University of Pelotas, Brazil

ARTICLE INFO

Keywords:
Physical activity
Accelerometry
Calibration study
Raw accelerometer data

ABSTRACT

Most of calibration studies based on accelerometry were developed using count-based analyses. In contrast, calibration studies based on raw acceleration signals are relatively recent and their evidences are incipient. The aim of the current study was to systematically review the literature in order to summarize methodological characteristics and results from raw data calibration studies. The review was conducted up to May 2017 using four databases: PubMed, Scopus, SPORTDiscus and Web of Science. Methodological quality of the included studies was evaluated using the Landis and Koch's guidelines. Initially, 1669 titles were identified and, after assessing titles, abstracts and full-articles, 20 studies were included. All studies were conducted in high-income countries, most of them with relatively small samples and specific population groups. Physical activity protocols were different among studies and the indirect calorimetry was the criterion measure mostly used. High mean values of sensitivity, specificity and accuracy from the intensity thresholds of cut-point-based studies were observed (93.7%, 91.9% and 95.8%, respectively). The most frequent statistical approach applied was machine learning-based modelling, in which the mean coefficient of determination was 0.70 to predict physical activity energy expenditure. Regarding the recognition of physical activity types, the mean values of accuracy for sedentary, household and locomotive activities were 82.9%, 55.4% and 89.7%, respectively. In conclusion, considering the construct of physical activity that each approach assesses, linear regression, machine-learning and cut-point-based approaches presented promising validity parameters.

1. Introduction

Questionnaires have historically been the main physical activity measurement instrument in epidemiological studies. However, accelerometers are currently a feasible alternative to objectively measure physical activity. Accelerometers are portable devices, which measure the acceleration from body movements in one, two or three axes: vertical (Y), horizontal right-left (X) and horizontal front-back axis (Z) [1].

The use of accelerometers entails advantages and disadvantages, as any other method of measurements of physical activity. Regarding the disadvantages, accelerometers do not indicate the context and the purpose of the physical activities. Furthermore, accelerometers are not valid to measure specific physical activities such as isometric activities, physical activities against a resistance force (e.g. strength exercises) and cycling [2]. However, the major source of overall physical activity energy expenditure (PAEE) is derived form dynamic physical activities (e.g. walking, running), and these activities are accurately measured by accelerometers [2]. Data from accelerometers are also free of information bias introduced by interviewers or participants. The data are gathered by the devices at the exact moment in which the activities are taking place, providing a reliable physical activity measure in free-living conditions [2,3].

An important challenge regarding the use of accelerometers to measure physical activity lies in the interpretation of the signals provided by the devices, which need to be translated into measurements with biological and/or behavioral meaning. In this context, several calibration studies have been performed [4]. There are important methodological differences in calibration studies (e.g. sample sizes and characteristics, physical activity protocols and statistical approaches), which might influence the results of such studies. Accordingly, the accelerometer signal, which is one of the main variables analysed in these studies is not the same across studies. Some studies, notoriously the most recent ones, have analysed the signal as a gravitational equivalent (g, where $1g = 9.81\ \mathrm{m\,s^{-2}}$), whilst other analysed it as

* Correspondence to: Rua Marechal Deodoro, 1160, CEP: 96020-220, Pelotas, Brazil.
E-mail addresses: marciopenha.esef@yahoo.com.br (M. de Almeida Mendes), inacio_cms@yahoo.com.br (I.C.M. da Silva), virgilioramires@hotmail.com (V.V. Ramires), ffreichert@gmail.com (F.F. Reichert), rafamartins1@gmail.com (R.C. Martins), tomasiet@gmail.com (E. Tomasi).

https://doi.org/10.1016/j.gaitpost.2017.12.028
Received 27 July 2017; Received in revised form 17 December 2017; Accepted 29 December 2017
0966-6362/

counts. Direct comparison between count values from different accelerometer brands is limited, mainly because manufacturers use different undisclosed algorithms to define the acceleration outputs [5]. In contrast, analyses based on gravitational equivalent (raw data) are performed using open source packages and, therefore, provides more transparency and better comparability across studies.

More recently accelerometer data can be analysed using different acceleration signals, thus, a high number of calibration studies based on raw data have been carried out. In this context, it is crucial to understand how the raw signal from accelerometers has been translated into physical activity measures. Thus, the aim of this study was to systematically review the literature in order to summarize the methods and results from calibration studies based on raw accelerometer data to measure physical activity.

2. Methods

The systematic review was conducted up to May 30th 2017 using PubMed, Scopus, SPORTDiscus and Web of Science databases. The following terms were searched in abstracts and titles: [("motor activity" OR "physical activity" OR "physical fitness" OR "physical exercise") AND ("accelerometry" OR "accelerometer" OR "motion sensor") AND ("calibration" OR "cut-off" OR "cut-point" OR "threshold" OR "validity" OR "validation") AND ("raw acceleration" OR "raw data")].

Only articles assessing raw acceleration signal and including healthy people were included. Articles in which the calibration was a secondary purpose were also eligible. Studies identifying only sedentary behavior thresholds were not considered.

Initially, all identified titles and abstracts were read by the first and third authors. In case of disagreement between them, all eligibility criteria were discussed. Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) guidelines were followed to conduct and describe all methodological process and reported results [6]. The following data were extracted from each study included in the review: place of publication, sample, accelerometer placement, accelerometer model, sample frequency and epoch (interval in which acceleration signals are summarized), activity protocol used for calibration, criterion measure, statistical approach applied, physical activity intensity thresholds and prediction equations of PAEE. Finally, the methodological quality of the included studies was evaluated using the Landis and Koch's guidelines [7].

3. Results

In total, 1669 articles were identified (23 references in PubMed, 19 in Scopus, 1590 in SPORTDiscus and 37 in Web of Science). After checking duplicate studies, 45 references were excluded and 1624 titles were considered eligible for reading. From the 1624, 180 titles were kept. After evaluating the abstracts, 79 articles remained. All 79 articles were read and 17 studies were considered eligible. The reference list from selected studies was checked and three articles were added, resulting in a total of 20 studies [8–27] (Fig. 1).

A detailed description of all studies included is available in Table 1. Studies were published between 1994 and 2017 and most of them were carried out in the United States of America (eight) and the remaining in a European country. Both sexes were included in all studies and there was no information regarding body composition and physical fitness in the studied samples. Eight studies were performed with children/adolescents and adults, eight exclusively with adults and three with children/adolescents. Regarding the studies with adults, only four of them presented age range greater than four decades. Studies with adolescents included participants from six to eighteen years old (Table 1).

The number of physical activities included in the protocols ranged from two to 23 and included a broad spectrum of intensity. Walking and running, as main components of PAEE, were included in 16 studies. Only four studies included physical activities practiced outside the laboratory setting (see Supplementary Table S1 in the online version at DOI:10.1016/j.gaitpost.2017.12.028).

Table 1 indicates that most studies (15) placed the accelerometer on the waist of the participants, but it was also tested on other parts of the body (low back, wrist, foot, chest, waist, thigh and ankle).

Actigraph GT3X (nine studies) and GENEActiv (five studies) were the main accelerometers used. Other accelerometers used were: ICSensors 3031–010, 7164 Actigraph, IDEEA, GENEA, Tracmor, DynaPort, Hookie AM13, Hookie AM20, GulfCoast X6-1A, MotionLogs and MICA2DOT (Table 1). Except for the 7164 Actigraph (1 axis), MICA2DOT and IDEEA (two axes), all other accelerometers collected body movements into three axes. The sampling frequency (number of measurements in each axis per second) varied from 10 to 100 Hz, and epoch lengths were analysed as one, five, six, 30 and 60 s (Table 1).

Regarding the studies in which the three-dimensional raw data were transformed into a single-dimensional signal vector magnitude (SVM) of acceleration, this conversion was performed using different metrics and the equation $SVM = \Sigma\sqrt{x^2+y^2+z^2}$ is the most common metric adopted. Indirect calorimetry was the most widely used criterion measure (13 studies) (Table 1).

Among the studies using cut-point-based statistical approach, the mean values and standard deviation (± SD) of sensitivity, specificity and accuracy from the intensity thresholds were: 93.7% (± 7.0), 91.9% (± 9.6) and 95.8% (± 0.1), respectively. Values of sensitivity, specificity or accuracy were similar according to the different intensity thresholds and accelerometer placements. The values of accuracy ranged from 84% to 100% (Table 2).

Five studies used regression models to estimate PAEE. Four of these studies presented predictive equations, in which the mean value of coefficient of determination (R^2) was 0.79 (± 0.12) (Table 3).

Machine learning-based modelling to estimate PAEE and to recognize physical activity types was the calibration statistical approach most frequently applied (11 studies). Most of these studies used Artificial Neural Network technique to create their predictive models (seven studies). Among the predictive models for PAEE, the mean value of R^2 was 0.70 (± 0.11). Regarding predictive models for recognition of physical activity types, the mean values of accuracy for sedentary (e.g. lying down, sitting, standing), household (washing dishes, folding towels and stacking them nearby, vacuuming carpet) and locomotive (walking, cycling, running) activities were 82.9% (± 20.2), 55.4% (± 26.6) and 89.7% (± 11.2), respectively (Table 4).

4. Discussion

All articles found in this review were conducted in high-income countries and most of them had relatively small samples and specific population groups, with low variability in terms of individual characteristics. The sample composition from accelerometry calibration studies hinders the extrapolation of their results to other settings [4]. Greater heterogeneity is required regarding the characteristics such as age, body mass index and physical fitness. Future calibration studies using more representative samples of their target populations in terms of demographic and physiological characteristics would represent an important step forward.

Distinct physical activity protocols were applied in the included calibration studies and were summarized in the present review. Protocols including the whole spectrum of physical activity intensities (sedentary pursuits, low, moderate and vigorous activities) were identified and the number of physical activities performed varied across studies. It is important to highlight that the number of activities tested shall not affect the internal validity of the intensity thresholds or algorithms. Therefore, high accuracy in the prediction of PAEE might be found even in studies assessing few activities. In contrast, a low number of activities tested or the inclusion of activities that are rarely performed in free-living conditions by the target population could impair