



Measuring Physical Activity in Free-Living Conditions—Comparison of Three Accelerometry-Based Methods

Anna-Maiju Leinonen^{1,2,3*}, Riikka Ahola^{1,4}, Janne Kulmala⁵, Harto Hakonen⁵, Henri Vähä-Ypyä⁶, Karl-Heinz Herzig^{4,7,8,9}, Juha Auvinen^{4,10}, Sirkka Keinänen-Kiukaanniemi^{4,10}, Harri Sievänen⁶, Tuija H. Tammelin⁵, Raija Korpelainen^{3,4,10} and Timo Jämsä^{1,2,4,11}

OPEN ACCESS

Edited by:

Gary Iwamoto,
University of Illinois at
Urbana–Champaign, USA

Reviewed by:

Edward Mitchell Winter,
Sheffield Hallam University, UK
Juliano Dal Pupo,
Universidade Federal de Santa
Catarina, Brazil
Weimo Zhu,
University of Illinois at
Urbana–Champaign, USA

*Correspondence:

Anna-Maiju Leinonen
anna.jauho@oulu.fi

Specialty section:

This article was submitted to
Exercise Physiology,
a section of the journal
Frontiers in Physiology

Received: 20 July 2016

Accepted: 22 December 2016

Published: 10 January 2017

Citation:

Leinonen A-M, Ahola R, Kulmala J, Hakonen H, Vähä-Ypyä H, Herzig K-H, Auvinen J, Keinänen-Kiukaanniemi S, Sievänen H, Tammelin TH, Korpelainen R and Jämsä T (2017) Measuring Physical Activity in Free-Living Conditions—Comparison of Three Accelerometry-Based Methods. *Front. Physiol.* 7:681. doi: 10.3389/fphys.2016.00681

¹ Research Unit of Medical Imaging, Physics and Technology, University of Oulu, Oulu, Finland, ² Infotech Oulu, University of Oulu, Oulu, Finland, ³ Department of Sports and Exercise Medicine, Oulu Deaconess Institute, Oulu, Finland, ⁴ Medical Research Center, Oulu University Hospital and University of Oulu, Oulu, Finland, ⁵ LIKES - Research Center for Sport and Health Sciences, Jyväskylä, Finland, ⁶ UKK Institute for Health Promotion Research, Tampere, Finland, ⁷ Research Unit of Biomedicine, University of Oulu, Oulu, Finland, ⁸ Department of Gastroenterology and Metabolism, Poznan University of Medical Sciences, Poznan, Poland, ⁹ Biocenter Oulu, University of Oulu, Oulu, Finland, ¹⁰ Center for Life Course Health Research, University of Oulu, Oulu, Finland, ¹¹ Diagnostic Radiology, Oulu University Hospital, Oulu, Finland

We examined the agreement in time spent on different physical activity (PA) levels using (1) mean amplitude deviation (MAD) of raw acceleration from the hip, (2) wrist-worn Polar Active, and (3) hip-worn Actigraph counts using Freedson's cut-points among adults under free-living conditions. PA was measured in 41 volunteers (mean age 47.6 years) for 14 days. Two MET-based threshold sets were used for MAD and Polar Active for sedentary time (ST) and time spent in light (LPA), moderate (MPA), and vigorous (VPA) PA. Actigraph counts were divided into PA classes, ≤ 100 counts/min for ST and Freedson's cut-points for LPA, MPA, and VPA. Analysis criteria were simultaneous use of devices for at least 4 days of >500 min/d. The between-method differences were analyzed using a repeated measures analysis of variance test. Bland-Altman plots and ROC graphs were also employed. Valid data were available from 27 participants. Polar Active produced the highest amount of VPA with both thresholds (≥ 5 and ≥ 6 MET; mean difference 17.9–30.9 min/d, $P < 0.001$). With the threshold 3–6 MET for MPA, Polar Active indicated 19.2 min/d more than MAD (95% CI 5.8–32.6) and 51.0 min/d more than Actigraph (95% CI 36.7–65.2). The results did not differ with 3.5–5 MET for MPA [$F_{(1.44, 37.43)} = 1.92$, $P = 0.170$]. MAD and Actigraph were closest to each other for ST with the threshold <1.5 MET (mean difference 22.2 min/d, 95% CI 7.1–37.3). With the threshold <2 MET, Polar Active and Actigraph provided similar results (mean difference 7.0 min/d, 95% CI –17.8–31.7). Moderate to high agreement (area under the ROC curve 0.806–0.963) was found between the methods for the fulfillment of the recommendation for daily moderate-to-vigorous PA of 60 min. In free-living conditions the agreement between MAD, Polar Active, and Actigraph for measuring time spent on different activity levels in adults was dependent on the activity thresholds used and PA intensity. ROC

analyses showed moderate to high agreement for the fulfillment of the recommendation for daily MVPA. Without additional statistical adjustment, these methods cannot be used interchangeably when measuring daily PA, but any of the methods can be used to identify persons with insufficient daily amount of MVPA.

Keywords: accelerometer, agreement, middle-aged, objective measurement, sedentary time

INTRODUCTION

Regular physical activity (PA) reduces the risk of many diseases, such as cardiovascular diseases, metabolic syndromes, type 2 diabetes, certain cancers, and depression (McGuire and Ross, 2011; Wen et al., 2011; Ekelund et al., 2012; Lee et al., 2012; Sieverdes et al., 2012). On the contrary, sedentary behavior has recently been suggested to pose even greater health risk than obesity (Ekelund et al., 2015). Despite the growing evidence for the positive health effects of PA, our lifestyle is becoming more inactive worldwide (Hallal et al., 2012). To improve specification of the amount of health enhancing PA, objective methods to measure PA, including sedentary time, are needed.

Nowadays, several types of accelerometers and analysis methods to measure the intensity, frequency, and duration of PA and sedentary time are available. The challenge is that the output of different methods varies and there still exists lack of consensus which of the methods provides the accurate results (Welk et al., 2012). Based on studies in laboratory, in short-term free-living conditions or under structured protocols, the output seems to be related to the used algorithms and thresholds of different PA levels (Strath et al., 2003; Reilly et al., 2008; Trost et al., 2011; Crouter et al., 2013; Thiese et al., 2014; Watson et al., 2014). Recently, the use of the accelerometers that provide raw acceleration signal instead of some specific unit (e.g., counts) has increased. However, transparent algorithms to analyse raw acceleration signal are still missing (Troiano et al., 2014).

Mean amplitude deviation (MAD) is a recently proposed universal method to analyse raw triaxial acceleration signals (Aittasalo et al., 2015; Vähä-Ypyä et al., 2015a,b). A strong correlation ($R^2 = 0.94$) has been found between incident MAD measured with a waist-worn device (Hookie AM20, Traxmeet, Finland) and oxygen consumption in walking and running (Vähä-Ypyä et al., 2015a). For the epoch of interest, the MAD is calculated from the resultant of the three orthogonal accelerations and it describes the mean variation of the dynamic acceleration component around the static component. On the other hand, Polar Active represents wrist-worn accelerometer with high user-friendliness (Brugniaux et al., 2010; Schaefer et al., 2014; Troiano et al., 2014; Jauho et al., 2015). In earlier studies Polar Active has been found to relate to indirect calorimetry and double labeled water with regard to energy expenditure (EE) calculation (Brugniaux et al., 2010; Kinnunen et al., 2012). Actigraph is currently the most widely used accelerometer for research purposes (Robusto and Trost, 2012; Santos-Lozano et al., 2012; Troiano et al., 2014). In previous studies, several Actigraph prediction equations for estimating EE have been developed (Crouter et al., 2006).

Although, there are some studies on convergent validity between accelerometry-based methods used in research (Welk et al., 2000; Paul et al., 2007; Esliger et al., 2011; John et al., 2013; Hildebrand et al., 2014; Lee et al., 2015; Rowlands et al., 2015), further equivalency studies are needed to better cover the wide range of available measurement methods. The above-mentioned three different accelerometer-based PA measurement methods have not been compared so far. Thus, the present study aims to produce this missing knowledge by comparing these methods, including one method developed to analyze raw acceleration data. Accordingly, the aim of this study was to examine the convergent validity in measured time spent on different PA levels calculated with MAD using raw acceleration data, Polar Active, and the Freedson equation for Actigraph counts among adults in free-living conditions.

MATERIALS AND METHODS

Participants

The study is a part of the pilot study of 46-year data collection of The Northern Finland Birth Cohort 1966 (NFBC1966) (Rantakallio, 1988). For the pilot study, 150 invitations were sent to adults selected randomly from the national population register, born in 1964–65 and living in Oulu area or in neighboring municipalities. A total of 45 respondents were recruited, their mean age being 47.6 (SD 0.6) years. In the beginning, the study participants were invited to a baseline visit, where their body mass and height were measured and they received the accelerometers with oral and written instructions. In addition, the participants received a prepaid-postage padded envelope for returning the devices by mail.

Of the 45 participants recruited, 41 (91%) attended the baseline visit and agreed to use the three accelerometers for 2 weeks. Simultaneous activity data comprising at least four valid days from all three devices were available from 29 (71%) participants. Three participants wore the monitors for 1 day only, in four cases the battery of one of the monitors had discharged, and in five cases one of the monitors had malfunctioned and the data were not recorded. During the analyses, two cases with invalid data were observed and were excluded from further analyses. In the end, 27 (66%) participants were included in the analyses. The characteristics of the 27 participants are shown in **Table 1**.

All participants were asked for a written consent to take part in the study. Prior to obtaining consent, the participants were given both oral and written information about the study and their right to refuse to take part or to withdraw from the study. The study followed the legislation decrees and ethical principles concerning medical research on humans in Finland and the Declaration of

TABLE 1 | Characteristics of the study participants included in the analyses.

	Females (<i>n</i> = 20)	Males (<i>n</i> = 7)
Age [years]	47.6 (0.6) [47–49]	47.7 (0.7) [47–49]
Height [cm]	164.6 (6.3) [150.8–176.0]	177.7 (6.7) [165.1–189.0]
Body mass [kg]	72.6 (13.6) [51.6–102.3]	80.9 (8.3) [65.1–89.5]
BMI [kg/m ²]	26.7 (4.6) [18.9–38.6]	25.6 (2.4) [20.9–28.5]

Values are mean (SD) [range].

Helsinki. The permission for the NFBC1966 study was received from the Ethics Committee of Northern Ostrobothnia Hospital District.

Objective Measurements of PA

The PA of the study participants was measured using three activity monitors (Hookie AM20, Polar Active, and Actigraph GT3X) continuously during a 2-week period. The activity measurement lasted for 15 days (including the day when the monitors were given). The monitors gave no feedback to the user.

Hookie AM20 (6.6 × 2.7 × 1.3 cm, mass 15 g, measurement range ± 16 g, Traxmeat Ltd., Espoo, Finland) is a triaxial accelerometer which can be worn on the hip on an elastic belt. Hookie AM20 collects raw acceleration data with a sampling frequency of 100 Hz. In this study, the epoch length was 30 s. For each epoch, MAD of the resultant acceleration was calculated from the raw acceleration data and converted to metabolic equivalents (METs) (Vähä-Ÿpyä et al., 2015a,b).

Polar Active (mass 45 g, frequency range 0.3–4.0 Hz, Polar Electro Ltd., Kempele, Finland) is a waterproof, wristwatch-like uniaxial accelerometer with a 21-day memory. The participants were advised to wear the device on the non-dominant hand for 24 h a day (except in the sauna bath). Using sex, age, body mass, and height as an input, Polar Active converts the measured acceleration signal to METs with the epoch length of 30 s. Polar Active has been shown to correlate well ($R^2 = 0.74$) with the doubly labeled water technique while assessing EE during a 7-day period of military training (Kinnunen et al., 2012). A high correlation ($R^2 = 0.97$) has also been obtained between EE measured with Polar Active prototype and indirect calorimetry during a 9.7-km hike (Brugniaux et al., 2010).

Actigraph GT3X (3.8 × 3.7 × 1.8 cm, mass 27 g, measurement range 0.05–2.5 g, Actigraph Inc., Pensacola, Florida, USA) is a hip-worn triaxial accelerometer which provides counts with the sampling frequency of 30 Hz. Actigraph has been widely used in large epidemiological studies (Robusto and Trost, 2012; Santos-Lozano et al., 2012; Troiano et al., 2014). During the data collection, the epoch length for Actigraph was 10 s which was transformed to 30 s for further analysis. Only the data from the vertical axis were included in the analysis. Previous studies have shown consistent results between vertical and triaxial Actigraph measurements (Kelly et al., 2013). Actigraph was worn on the same belt with the Hookie accelerometer. The participants were advised to wear the belt with both hip-worn activity monitors being on the right side during the waking hours, excluding the time spent in the sauna, bath, shower, or in other water activities.

The three PA measurement methods were compared within the time periods when all three devices had been simultaneously used. To be eligible for the analyses, all accelerometers had to be used simultaneously for at least four valid days; a valid day being at least 500 min of wearing time per day. The non-wearing time was defined as at least a 60-min period of consecutive zeros. Total between-subject daily time (in minutes) spent in sedentary time (ST), light (LPA), moderate (MPA), and vigorous (VPA) PA were compared.

For calculating the daily time spent on different PA levels, two different sets of MET-based thresholds were used for both MAD and Polar Active:

(SET 1) ST: ≤1.5 METs, LPA: 1.51–2.99 METs, MPA: 3–5.99 METs, and VPA: ≥6 METs; and

(SET 2) ST: <2 METs, LPA: 2–3.49 METs, MPA: 3.5–4.99 METs, and VPA: ≥5 METs.

The SET 1 is widely used (Plasqui et al., 2013; Gibbs et al., 2015) whereas the SET 2 is standard in Polar Active software and analyses. Actigraph counts were divided into MET-based PA classes using ≤100 counts/min for ST (Matthews et al., 2008) and the Freedson thresholds for LPA (101–1951 counts/min), MPA (1952–5724 counts/min), and VPA (≥5725 counts/min) (Freedson et al., 1998). The personal mean values over the period of valid data for each PA measurement method and activity level were calculated for the further analyses.

Statistical Analysis

The results were analyzed with IBM SPSS Statistics software (SPSS 19 for Windows, SPSS Inc., Chicago, Illinois). A *P*-value below 0.05 was considered statistically significant. At the method level, the mean time spent on different activity levels was calculated through personal mean values, and the between-method differences in ST, LPA, MPA, and VPA were analyzed using a repeated measures analysis of variance (ANOVA) test. If the assumption of sphericity was violated, the results with a Greenhouse-Geisser correction was presented. A post hoc test with Bonferroni adjustment for multiple comparisons was used. Cohen's *d*-values were used to determine the effect sizes.

Intra-class correlation coefficients (ICC) and Bland-Altman method were used for assessing agreement between the different methods at the individual level. In these analyses, the active time was defined as the sum of MPA and VPA, i.e., moderate-to-vigorous PA (MVPA). The correlation coefficient indicates the strength of association between two methods but the actual agreement is illustrated by the Bland-Altman method (Bland and Altman, 1986). The mean difference, standard deviation of the difference and 95% limits of agreement were calculated.

Receiver operating characteristics (ROC) graphs were used to evaluate and compare the used methods in classifying daily time spent in MVPA (Fawcett, 2006). Both MET thresholds for MVPA time, 3 and 3.5 METs, were included to the analysis. Each of the three methods was in turn used as a reference. The classification of a reference method was based on the fulfillment of PA recommendation which in this case was selected to be 60 min of MVPA time per day (Hallal et al., 2012). To estimate the

discriminatory accuracy of each method, the area under the ROC curve (AUC) was calculated.

RESULTS

As a total, 376 valid days were included to the analyses. From these days, 98% exceeded also the limit of 600 min of wearing time per day. The mean number of valid days was 14 (SD 2.0) per person, and the mean wear time was 15.1 (SD 1.4) h/d. An example of 1-day activity data calculated to MAD (Hookie), METs (Polar Active), and Actigraph counts is presented in **Figure 1**. The daily patterns of activity shown as MAD and counts were rather similar. The MET values produced by Polar Active occasionally had the highest activity peaks at different time points compared to hip-worn devices.

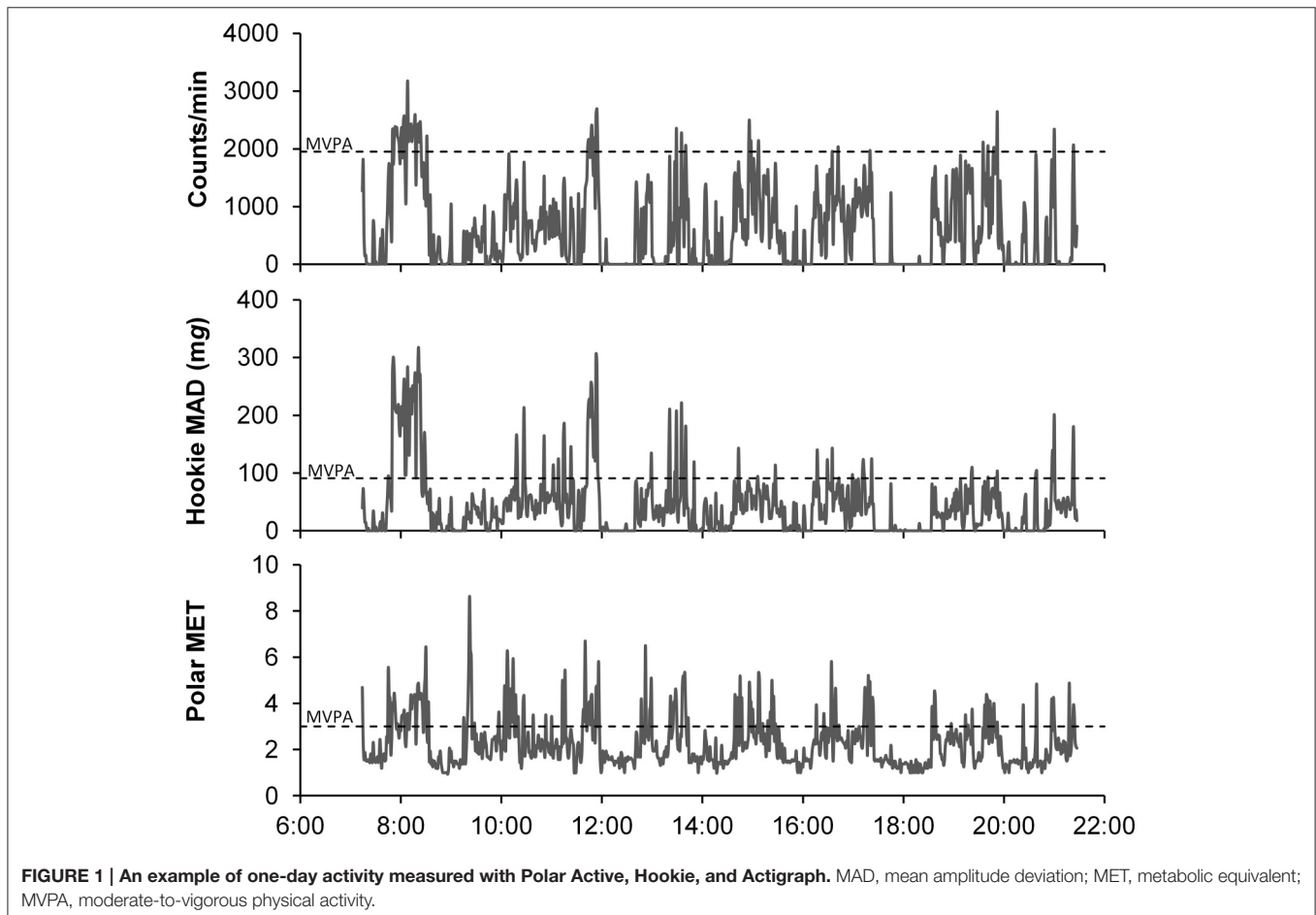
The mean values for all four activity levels are shown in **Figure 2**, including the results based on the Freedson thresholds for Actigraph and both MET threshold sets (*SETs 1 and 2*) for MAD and Polar Active. When ST was analyzed with threshold <1.5 MET, values produced by MAD and Actigraph were closest to each other but yet differed significantly (mean difference 22.2 min/d, 95% CI 7.1–37.3, $P < 0.01$, Cohen's $d = 0.238$). Instead, with the threshold <2 MET Polar Active and Actigraph gave

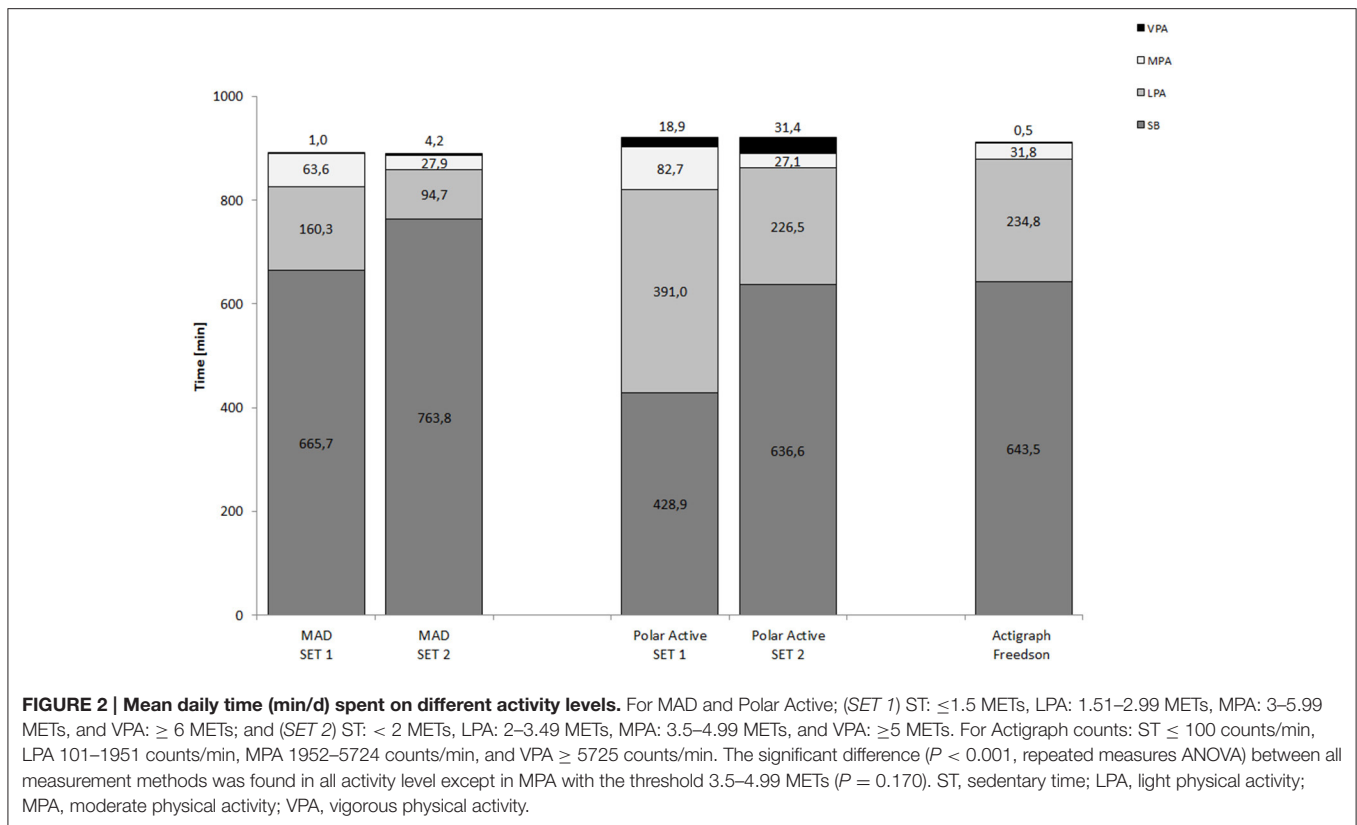
similar results (mean difference 7.0 min/d, 95% CI 17.8–31.7, $P = 1.0$, $d = 0.064$).

Compared to the other two methods, the mean LPA time was the smallest analyzed by MAD with both thresholds (mean difference 74.5–230.7 min/d, $P < 0.001$, $d = 1.399$ –3.726). The results of Polar Active and Actigraph were close to each other with the threshold 2–3.5 MET (mean difference 8.2 min/d, 95% CI 12.9–29.4, $P = 0.985$, $d = 0.123$). Instead, with the threshold 1.5–3 MET Polar Active resulted significantly higher amount of LPA compared to Actigraph (mean difference 156.2 min/d, 95% CI 131.9–180.5, $P < 0.001$, $d = 2.315$).

When the threshold 3–6 MET for MPA time was used, the value indicated by Polar Active was the highest, being 19.2 min/d higher than MAD (95% CI 5.8–32.6, $P < 0.01$, $d = 0.733$) and 51.0 min/d higher than Actigraph (95% CI 36.7–65.2, $P < 0.001$, $d = 2.175$). Using the threshold 3.5–5 MET, the results were close to each other with all three accelerometers [$F_{(1.44, 37.43)} = 1.92$, $P = 0.170$].

The mean VPA time was the highest measured with Polar Active compared to MAD and Actigraph with both threshold sets (mean difference 17.9–30.9 min/d, $P < 0.001$, $d = 1.313$ –2.020). For Polar Active the difference in results between the threshold *SET 1* and *2* was 13 min/d ($d = 1.313$). The results for VPA were close to zero produced by MAD with the threshold >6 MET





and by Actigraph with the Freedson threshold (1.0 vs. 0.5 min/d, $P = 0.526$, $d = 0.335$). With the threshold > 5 MET MAD resulted in 4 min difference to Actigraph (95% CI 0.8–6.5, $P < 0.01$, $d = 0.884$).

The Bland-Altman plots illustrating the agreement between different methods in estimating time spent on different activity levels (analyzed with the threshold SET 1) are shown in **Figure 3**. The highest agreement was found between MAD and Actigraph on ST (ICC = 0.922). For MVPA with the threshold > 3 MET, all measurement methods gave divergent results. Polar Active overestimated time spent in MVPA by 37.1 min/d (SD 31.6 min/d) compared to MAD and by 69.3 min/d (SD 37.7 min/d) compared to Actigraph.

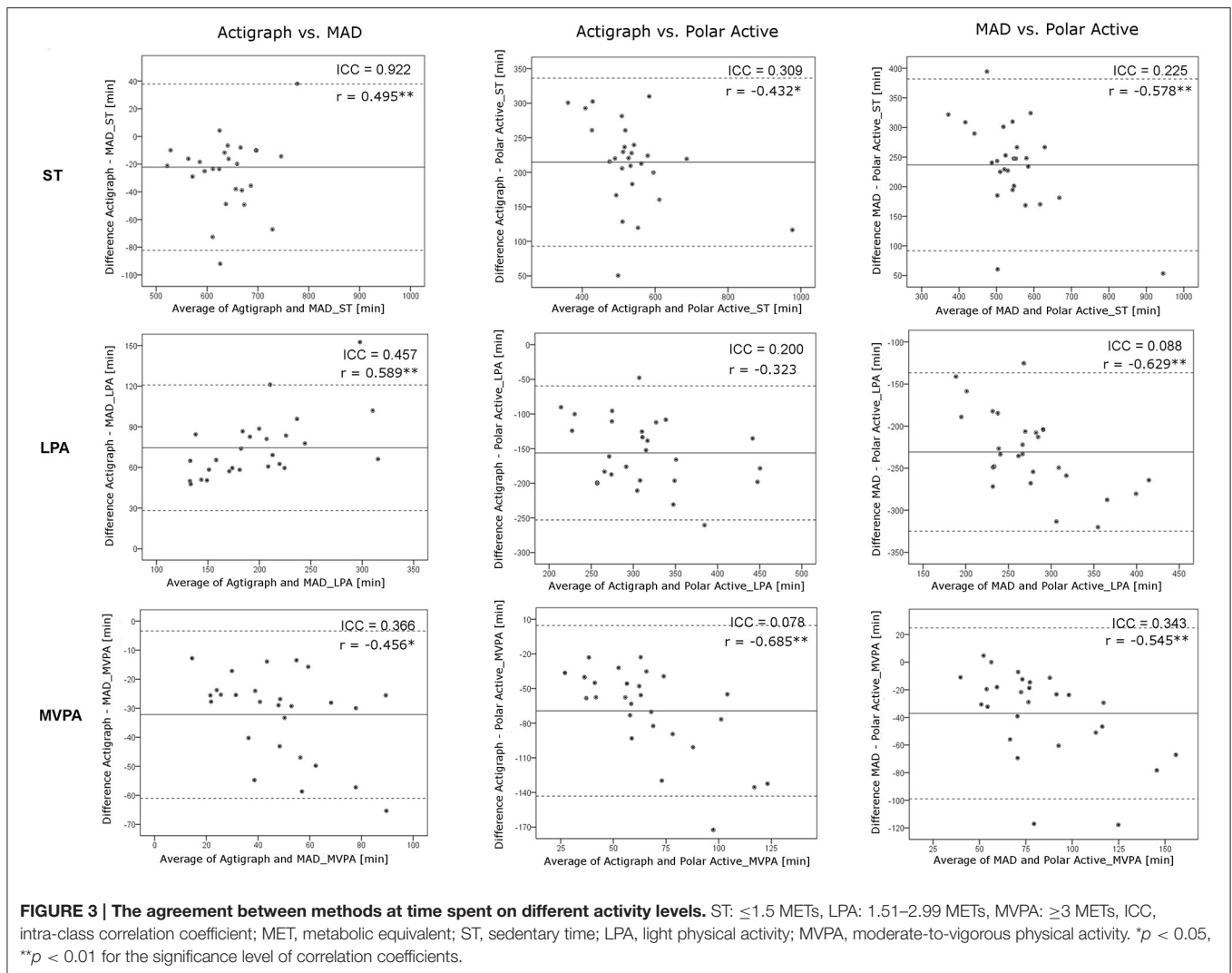
The ROC curves for comparisons of the fulfillment of MVPA recommendation measured by MAD, Polar Active, and Actigraph with the Freedson thresholds are shown in **Figure 4**. Moderate to high agreement between all the methods was found, the best degree of consistency observed with the threshold value of 3.5 MET, Actigraph being the reference method [AUC for MAD = 0.963 (95% CI 0.934–0.991), AUC for Polar Active = 0.895 (95% CI 0.860–0.929)].

DISCUSSION

In this study, the agreement in time spent on different PA intensity levels calculated using MAD for triaxial raw acceleration data, Polar Active, and the Freedson equation for Actigraph counts was examined among middle-aged people. The findings

showed that in daily life conditions the agreement between the used methods varied by the selected activity thresholds as well as by the intensity of PA. However, according to the ROC analyses, there is moderate to high agreement for the fulfillment of the recommendation for daily MVPA of 60 min.

Polar Active resulted in the lowest amount of ST and the highest amount of MVPA with both MET threshold sets for differentiating between different PA levels. The results of Polar Active were closer to the results of the two other methods when the standard Polar Active thresholds were used instead of the widely used threshold SET 1 (**Figures 2, 4**). The level of overestimation increased in wrist-worn Polar Active compared to the two hip-worn methods with higher values of accumulated MVPA time (significant magnitude-associations in Bland-Altman plots in **Figure 3**). The difference between the methods in MVPA time was most evident for the vigorous activity. The difference was also revealed in the daily pattern of activity (**Figure 1**). These results are in accordance with the previous findings stating that wrist-worn monitors show higher activity output than monitors worn on hip (Hildebrand et al., 2014). When indirect calorimetry has been used as the criterion measure, hip-worn monitors have been shown to be more accurate in comparison to wrist-worn monitors (Esliger et al., 2011; Rosenberger et al., 2013). However, recent studies showed that the outcomes are very similar between hip and wrist when appropriate site-specific equations are used (Troost et al., 2014; Crouter et al., 2015).



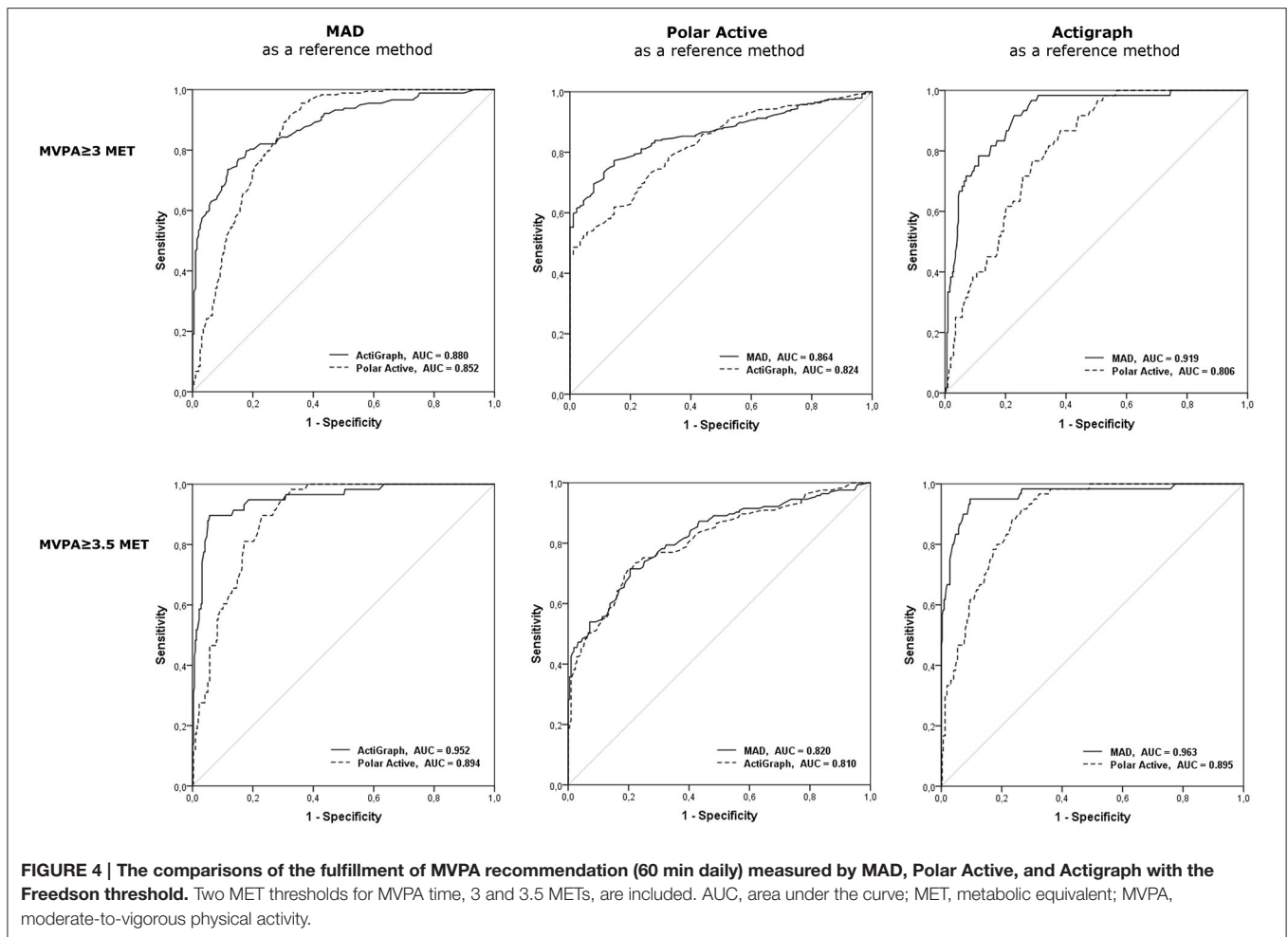
Deviating results in some examined activity levels were observed also between MAD and Actigraph based on the Freedson thresholds. MAD increasingly underestimated LPA but overestimated MVPA compared to Actigraph (significant magnitude-associations in Bland-Altman plots in **Figure 3**). However, comparable results between the devices might have been achieved if the raw Actigraph data had been analyzed with the universal MAD approach (Aittasalo et al., 2015; Vähä-Ypyä et al., 2015b). However, due to the long-term data collection, raw data were not available for the Actigraph device.

The thresholds used may exert even more effect on results than the monitor location (Rowlands et al., 2014). The Freedson thresholds have been stated to overestimate resting and LPA by 13% and underestimate MPA by 60%, compared to the results from the portable metabolic measurement system (Strath et al., 2003). These thresholds were developed based on treadmill walking and running which might be one reason for the difference in free-living conditions (Freedson et al., 1998). In the

present study the thresholds developed by Troiano et al., 2008 were additionally tested, resulting in similar values compared to the results analyzed with the Freedson thresholds (data not shown; Troiano et al., 2008).

The ROC analyses showed moderate to high agreement between the different methods for the fulfillment of the recommendation for daily MVPA of 60 min, which suggests that any of the methods can be used to identify persons with insufficient daily amount of MVPA. However, the used method and threshold affected the level of agreement (**Figure 4**).

In this study, we did not use indirect calorimetry or doubly labeled water technique as a reference method, which may be considered a limitation. However, these techniques cannot reveal specific patterns of PA or sedentary behavior but rather provide the mean EE over the period of interest. It is important to note that even if in some cases the consistency between devices was evident at a group level, differences at an individual level could still be quite large. Additionally, the sample presented a narrow age range of middle-aged men and women, limiting the



generalizability of the findings particularly to children or older adults.

The main strength of this study was the long 2-week measurement period in free-living environment. The monitors were conscientiously used by the study participants, the mean wear time satisfactorily covering the waking hours as well as the 2-week measurement period.

CONCLUSION

This study revealed that in daily life conditions the agreement between the MAD of raw triaxial acceleration, Polar Active, and Actigraph with the Freedson thresholds for measuring time spent on different activity levels in middle-aged people is highly dependent on the thresholds used as well as on the intensity level of physical activity. Good agreement between the methods was found for the fulfillment of the daily recommendation for MVPA of 60 min. However, based on the current results, it is not possible to say which of the methods is superior to each other. Instead, this study showed that without additional statistical adjustment, the results obtained from these types of accelerometers and analyses cannot be used interchangeably when estimating time spent on

different activity levels on daily life. Anyway, any of the methods can be used to identify persons with insufficient daily amount of MVPA.

AUTHOR CONTRIBUTIONS

RA, TJ, TT, RK, SKK, JA, KHH and AML have given significant contribution for the planning and implementation of the study, and for the planning and writing of the submitted manuscript. TJ, RK and SKK have been the principal investigators and have obtained funding for the study. JK and HH have been responsible for analyzing data from Polar Active and Actigraph. HVY and HS have been responsible for performing MAD analysis. The corresponding author AML has been responsible for statistical analysis and writing the first draft of the manuscript. All authors have revised, read, and accepted the final version of the manuscript.

ACKNOWLEDGMENTS

The authors acknowledge Dr. Paula Virtanen, Polar Electro for her support, and Mr. Risto Bloigu for statistical consultancy.

The authors would like to thank the volunteer study participants and the study nurses for their efforts in the field work. This work was financially supported by Infotech Oulu Doctoral Program, European Regional Development Fund

under Grant 539/2010 A31592, Ministry of Education and Culture, and Finnish Cultural Foundation. The financiers had no control on the conduct or the publication of the study.

REFERENCES

- Aittasalo, M., Vähä-Ypyä, H., Vasankari, T., Husu, P., Jussila, A., and Sievänen, H. (2015). Mean amplitude deviation calculated from raw acceleration data: a novel method for classifying the intensity of adolescents' physical activity irrespective of accelerometer brand. *BMC Sports Sci. Med. Rehabil.* 7:18. doi: 10.1186/s13102-015-0010-0
- Bland, J. M., and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1, 307–310. doi: 10.1016/S0140-6736(86)90837-8
- Brugniaux, J. V., Niva, A., Pulkkinen, I., Laukkanen, R. M., Richalet, J., and Pichon, A. P. (2010). Polar activity watch 200: a new device to accurately assess energy expenditure. *Br. J. Sports Med.* 44, 245–249. doi: 10.1136/bjsm.2007.045575
- Crouter, S. E., Churilla, J. R., and Bassett, D. R. Jr. (2006). Estimating energy expenditure using accelerometers. *Eur. J. Appl. Physiol.* 98, 601–612. doi: 10.1007/s00421-006-0307-5
- Crouter, S. E., DellaValle, D. M., Haas, J. D., Frongillo, E. A., and Bassett, D. R. (2013). Validity of actigraph 2-regression model, Matthews cut-points, and NHANES cut-points for assessing free-living physical activity. *J. Phys. Act. Health* 10, 504–514. doi: 10.1123/jpah.10.4.504
- Crouter, S. E., Flynn, J. I., and Bassett, D. R. (2015). Estimating physical activity in youth using a wrist accelerometer. *Med. Sci. Sports Exerc.* 47, 944–951. doi: 10.1249/MSS.0000000000000502
- Ekelund, U., Luan, J., Sherar, L. B., Esliger, D. W., Griew, P., and Cooper, A. (2012). Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. *J. Am. Med. Assoc.* 307, 704–712. doi: 10.1001/jama.2012.156
- Ekelund, U., Ward, H. A., Norat, T., Luan, J., May, A. M., Weiderpass, E., et al. (2015). Physical activity and all-cause mortality across levels of overall and abdominal adiposity in European men and women: the European Prospective Investigation into Cancer and Nutrition Study (EPIC). *Am. J. Clin. Nutr.* 101, 613–621. doi: 10.3945/ajcn.114.100065
- Esliger, D. W., Rowlands, A. V., Hurst, T. L., Catt, M., Murray, P., and Eston, R. G. (2011). Validation of the GENEA accelerometer. *Med. Sci. Sports Exerc.* 43, 1085–1093. doi: 10.1249/MSS.0b013e31820513be
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recogn. Lett.* 27, 861–874. doi: 10.1016/j.patrec.2005.10.010
- Freedson, P. S., Melanson, E., and Sirard, J. (1998). Calibration of the Computer Science and Applications, Inc. accelerometer. *Med. Sci. Sports Exerc.* 30, 777–781. doi: 10.1097/00005768-199805000-00021
- Gibbs, B. B., Hergenroeder, A. L., Katzmarzyk, P. T., Lee, I. M., and Jakicic, J. M. (2015). Definition, measurement, and health risks associated with sedentary behavior. *Med. Sci. Sports Exerc.* 47, 1295–1300. doi: 10.1249/MSS.0000000000000517
- Hallal, P. C., Andersen, L. B., Bull, F. C., Guthold, R., Haskell, W., Ekelund, U., et al. (2012). Global physical activity levels: surveillance progress, pitfalls, and prospects. *Lancet* 380, 247–257. doi: 10.1016/S0140-6736(12)60646-1
- Hildebrand, M., Van Hees, V. T., Hansen, B. H., and Ekelund, U. (2014). Age-group comparability of raw accelerometer output from wrist- and hip-worn monitors. *Med. Sci. Sports Exerc.* 46, 1816–1824. doi: 10.1249/MSS.0000000000000289
- Jauho, A., Pyky, R., Ahola, R., Kangas, M., Virtanen, P., Korpelainen, R., et al. (2015). Effect of wrist-worn activity monitor feedback on physical activity behavior: a randomized controlled trial in Finnish young men. *Prevent. Med. Reports* 2, 628–634. doi: 10.1016/j.pmedr.2015.07.005
- John, D., Sasaki, J., Staudenmayer, J., Mavilia, M., and Freedson, P. S. (2013). Comparison of raw acceleration from the genea and actigraph™ GT3X+ activity monitors. *Sensors* 13, 14754–14763. doi: 10.3390/s131114754
- Kelly, L. A., McMillan, D. G., Anderson, A., Fippinger, M., Fillerup, G., and Rider, J. (2013). Validity of Actigraphs uniaxial and triaxial accelerometers for assessment of physical activity in adults in laboratory conditions. *BMC Med. Phys.* 13:5. doi: 10.1186/1756-6649-13-5
- Kinnunen, H., Tanskanen, M., Kyröläinen, H., and Westerterp, K. R. (2012). Wrist-worn accelerometers in assessment of energy expenditure during intensive training. *Physiol. Meas.* 1841–1854. doi: 10.1088/0967-3334/33/11/1841
- Lee, I. M., Shiroma, E. J., Lobelo, F., Puska, P., Blair, S. N., Katzmarzyk, P. T., et al. (2012). Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *Lancet* 380, 219–229. doi: 10.1016/S0140-6736(12)61031-9
- Lee, J. A., Williams, S. M., Brown, D. D., and Laurson, K. R. (2015). Concurrent validation of the Actigraph gt3x+, Polar Active accelerometer, Omron HJ-720 and Yamax Digiwalker SW-701 pedometer step counts in lab-based and free-living settings. *J. Sports Sci.* 33, 991–1000. doi: 10.1080/02640414.2014.981848
- Matthews, C. E., Chen, K. Y., Freedson, P. S., Buchowski, M. S., Beech, B. M., Pate, R. R., et al. (2008). Amount of time spent in sedentary behaviors in the United States, 2003–2004. *Am. J. Epidemiol.* 167, 875–881. doi: 10.1093/aje/kwm390
- McGuire, K. A., Ross, R. (2011). Incidental physical activity is positively associated with cardiorespiratory fitness. *Med. Sci. Sports Exerc.* 43, 2189–2194. doi: 10.1249/MSS.0b013e31821e4ff2
- Paul, D. R., Kramer, M., Moshfegh, A. J., Baer, D. J., and Rumpler, W. V. (2007). Comparison of two different physical activity monitors. *BMC Med. Res. Methodol.* 7:26. doi: 10.1186/1471-2288-7-26
- Plasqui, G., Bonomi, A. G., and Westerterp, K. R. (2013). Daily physical activity assessment with accelerometers: new insights and validation studies. *Obes. Rev.* 14, 451–462. doi: 10.1111/obr.12021
- Rantakallio, P. (1988). The longitudinal study of the Northern Finland birth cohort of 1966. *Paediatr. Perinat. Epidemiol.* 2, 59–88. doi: 10.1111/j.1365-3016.1988.tb00180.x
- Reilly, J. J., Penpraze, V., Hislop, J., Davies, G., Grant, S., and Paton, J. Y. (2008). Objective measurement of physical activity and sedentary behaviour: review with new data. *Arch. Dis. Child.* 93, 614–619. doi: 10.1136/adc.2007.133272
- Robusto, K. M., and Trost, S. G. (2012). Comparison of three generations of ActiGraph™ activity monitors in children and adolescents. *J. Sports Sci.* 30, 1429–1435. doi: 10.1080/02640414.2012.710761
- Rosenberger, M. E., Haskell, W. L., Albin, F., Mota, S., Nawyn, J., and Intille, S. (2013). Estimating activity and sedentary behavior from an accelerometer on the hip or wrist. *Med. Sci. Sports Exerc.* 45, 964–975. doi: 10.1249/MSS.0b013e31827f0d9c
- Rowlands, A. V., Frayssé, F., Catt, M., Stiles, V. H., Stanley, R. M., Eston, R. G., et al. (2015). Comparability of measured acceleration from accelerometer-based activity monitors. *Med. Sci. Sports Exerc.* 47, 201–210. doi: 10.1249/MSS.0000000000000394
- Rowlands, A. V., Rennie, K., Kozarski, R., Stanley, R. M., Eston, R. G., Parfitt, G. C., et al. (2014). Children's physical activity assessed with wrist- and hip-worn accelerometers. *Med. Sci. Sports Exerc.* 46, 2308–2316. doi: 10.1249/MSS.0000000000000365
- Santos-Lozano, A., Marin, P. J., Torres-Luque, G., Ruiz, J. R., Lucía, A., and Garatachea, N. (2012). Technical variability of the GT3X accelerometer. *Med. Eng. Phys.* 34, 787–790. doi: 10.1016/j.medengphy.2012.02.005
- Schaefer, S. E., Van Loan, M., and German, J. B. (2014). A feasibility study of wearable activity monitors for pre-adolescent school-age children. *Prev. Chronic Dis.* 11:130262. doi: 10.5888/pcd11.130262
- Sieverdes, J. C., Ray, B. M., Sui, X., Lee, D., Hand, G. A., Baruth, M., et al. (2012). Association between leisure time physical activity and depressive symptoms in men. *Med. Sci. Sports Exerc.* 44, 260–265. doi: 10.1249/MSS.0b013e31822e00a6
- Strath, S. J., Bassett, D. R. Jr., and Swartz, A. M. (2003). Comparison of MTI accelerometer cut-points for predicting time spent in physical activity. *Int. J. Sports Med.* 24, 298–303. doi: 10.1055/s-2003-39504
- Thiese, M. S., Hegmann, K. T., Behrens, T. K., Garg, A., and Porucznik, C. (2014). Important differences in accelerometer cut points for quantifying physical

- activity in a nested occupational cohort. *J. Exerc. Sports Orthop.* 1:12. doi: 10.15226/2374-6904/1/1/00102
- Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., and McDowell, M. (2008). Physical activity in the United States measured by accelerometer. *Med. Sci. Sports Exerc.* 40, 181–188. doi: 10.1249/mss.0b013e31815a51b3
- Troiano, R. P., McClain, J. J., Brychta, R. J., and Chen, K. Y. (2014). Evolution of accelerometer methods for physical activity research. *Br. J. Sports Med.* 48, 1019–1023. doi: 10.1136/bjsports-2014-093546
- Trost, S. G., Loprinzi, P. D., Moore, R., and Pfeiffer, K. A. (2011). Comparison of accelerometer cut points for predicting activity intensity in youth. *Med Sci Sports Exerc.* 43, 1360–1368. doi: 10.1249/MSS.0b013e318206476e
- Trost, S. G., Zheng, Y., and Wong, W. (2014). Machine learning for activity recognition: hip versus wrist data. *Physiol. Meas.* 35, 2183–2189. doi: 10.1088/0967-3334/35/11/2183
- Vähä-Ypyä, H., Vasankari, T., Husu, P., Mänttari, A., Vuorimaa, T., Suni, J., et al. (2015a). Validation of Cut-Points for Evaluating the Intensity of Physical Activity with Accelerometry-Based Mean Amplitude Deviation (MAD). *PLoS ONE* 10:e0134813. doi: 10.1371/journal.pone.0134813
- Vähä-Ypyä, H., Vasankari, T., Husu, P., Suni, J., and Sievänen, H. (2015b). A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. *Clin. Physiol. Funct. Imaging* 35, 64–70. doi: 10.1111/cpf.12127
- Watson, K. B., Carlson, S. A., Carroll, D. D., and Fulton, J. E. (2014). Comparison of accelerometer cut points to estimate physical activity in US adults. *J. Sports Sci.* 32, 660–669. doi: 10.1080/02640414.2013.847278
- Welk, G. J., Blair, S. N., Wood, K., Jones, S., and Thompson, R. W. (2000). A comparative evaluation of three accelerometry-based physical activity monitors. *Med. Sci. Sports Exerc.* 32, 489–497. doi: 10.1097/00005768-200009001-00008
- Welk, G. J., McClain, J., and Ainsworth, B. E. (2012). Protocols for evaluating equivalency of accelerometry-based activity monitors. *Med. Sci. Sports Exerc.* 44, 39–49. doi: 10.1249/MSS.0b013e3182399d8f
- Wen, C. P., Wai, J. P., Tsai, M. K., Yang, Y. C., Cheng, T. Y., Lee, M., et al. (2011). Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. *Lancet* 378, 1244–1253. doi: 10.1016/S0140-6736(11)60749-6

Conflict of Interest Statement: RA is currently (from November 2015) employed by Polar Electro Ltd. She was not involved in the company during data collection and analysis. The company had no role in the conduct of the study or decision to publish.

The other authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer WZ and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2017 Leinonen, Ahola, Kulmala, Hakonen, Vähä-Ypyä, Herzig, Auvinen, Keinänen-Kiukaanniemi, Sievänen, Tammelin, Korpelainen and Jämsä. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.