ResearchGate

A method to compare new and traditional accelerometry data in physical activity monitoring

Conference Paper · June 2010

DOI: 10.1109/WOWMOM.2010.5534986 · Source: DBLP

Vincent T Van Hees
Netherlands eScience Center

Marcelo Pias
University of Cambridge

Ulf Ekelund
Norwegian School of Sport Sciences (NIH)

Soren Brage
University of Cambridge

A method to compare new and traditional accelerometry data in physical activity monitoring

Vincent T. van Hees[1], Marcelo Pias[2], Salman Taherian[2], Ulf Ekelund[1], Søren Brage[1]

[1] Medical Research Council (MRC)
Epidemiology Unit
Institute of Metabolic Science
Box 285, Addenbrooke's Hospital
Cambridge CB2 0QQ, UK
vincent.vanhees@mrc-epid.cam.ac.uk

[2] Computer Laboratory
University of Cambridge
William Gates Building
15 JJ Thomson Avenue
Cambridge CB3 0FD, UK
marcelo.pias@cl.cam.ac.uk

***Abstract*— The accelerometer devices as traditionally used in the epidemiological field for physical activity monitoring (e.g. Actigraph, Actical, and RT3) provide manufacturer-dependent output values called *counts* that are computed by obscure and proprietary signal processing techniques. This lack of transparency poses a challenge for comparison of historical accelerometer data in counts with data collected using raw accelerometry in S.I. units - m/s^2. The purpose of this study was to develop a method that facilitates the compatibility between both methods through conversion of raw accelerometer output data collected with inertial acceleration sensors into Actigraph counts - the most widely used (*de facto standard*) device brand in epidemiological studies. The basics of the conversion algorithm were captured from the technical specifications of the Actigraph GT1M. Fine-tuning of the algorithm was achieved empirically under controlled conditions using a mechanical shaker device. A pilot evaluation was carried out through physical activity monitoring in free-living scenarios of 19 adult participants (age: 47 ± 11 yrs, BMI: 25.2 ± 4.1 kg·m^{-2}) wearing both devices. The results show that Actigraph counts estimated by the proposed method explain 94.2% of the variation in Actigraph counts ($p < 0.001$). The concordance correlation coefficient was 0.93 ($p < 0.05$). The sensitivity for classifying intensity ranged from 93.4% for light physical activity to 70.7% for moderate physical activity.**

Keywords: Physical activity monitoring, accelerometer signal processing, Conversion algorithm

I. Introduction

Our society is increasingly exposed to the problem of 'lifestyle diseases' including overweight and obesity which can lead to more serious clinical conditions such as diabetes, cardiovascular disease and some cancers. In the UK, almost two thirds of adults and a third of children are either overweight or obese [14]. The pressure, that chronic diseases put on families, the UK National Health System and society more broadly, is enormous with an estimated cost of £6 to 7.4 billion [9]. On current trends nearly 60 per cent of the UK population will be obese by 2050, with a £50 billion/year cost [9]. Obesity has also a negative impact on the environment with carbon emissions from additional food and its transport [7]. The core of the problem is two-fold: eating unhealthy food and lack of regular physical activity. The solution is more difficult to realise as it requires significant behavioural changes in people. Personalised prevention strategies and objective assessment of human physical activity using pervasive sensing technologies are on the critical path to reverse the trends in excess weight [12].
Accelerometry is the most commonly applied method for objective assessment of physical activity in epidemiological studies. In the past twenty-five years several generations of accelerometer devices have emerged [10,13,18]. Specifically, one brand, namely the Actigraph, has been extensively used in large-scale field studies [5,6,23,24,25] and has become the *de facto standard* device for objective physical activity monitoring. As newer measurement devices and methods (mostly proposed by computer scientists [26]) find their way into epidemiological studies it is important to ensure that the information collected is also compatible with traditional measures of physical activity so that historical analysis and comparisons can be made unambiguously.

Problem statement: To enable comparison between new and traditional physical activity (accelerometry) data.

Contribution: This paper introduces a method for the conversion of raw accelerometer data to the output signal of the Actigraph. Additionally, preliminary results are discussed from a pilot evaluation of conversion algorithm based on data collected in free-living volunteers.

II. Background

Accelerometry is the most commonly applied method for objective assessment of physical activity in epidemiological studies. Most traditional accelerometer devices only store a summary measure of the raw acceleration signal, termed a "count" [6]. A count is a unit aimed to be proportional to the average overall acceleration of the human body in a specified period of time, referred to as an "epoch" [4].

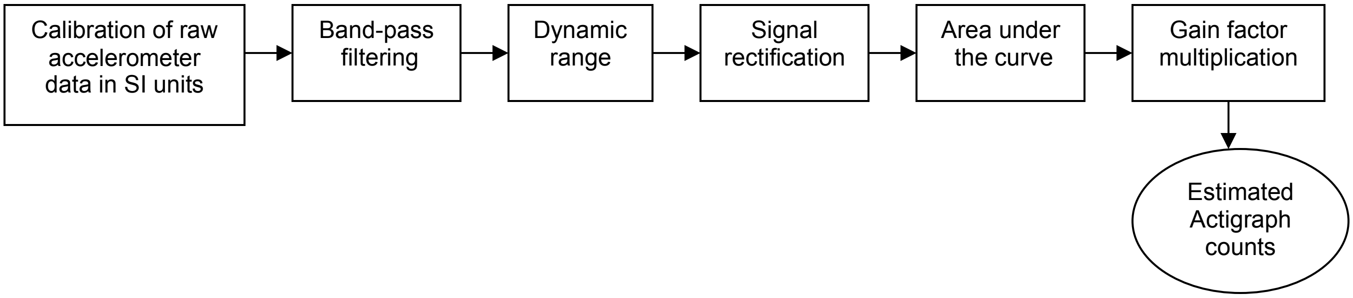

Figure 1. Processing blocks of the proposed mimicking method

However, this proportionality may be challenged by the dynamic range of the accelerometer and downstream signal processing, such as frequency filtering, implemented with the intention to remove components of the signal unrelated to human movements [18,2,15]. A variety of methods exist to filter and summarise the raw acceleration signal, the choice of which may have profound implications on the interpretation of the final output [16]. However, as traditional accelerometers are limited in memory and battery capacity to store raw signal data, such data processing stages are implemented on the device itself, and the process is irreversible once the count has been stored in local memory. This irreversible conversion prevents re-analysis of the raw accelerometer signal with the latest insights into data processing techniques.

Although the complete scientific description of a method would include sufficient detail of the signal processing scheme used in order to enable replication of empirical evidence, most manufacturers of accelerometer devices in use today regard the way in which the raw data is pre-processed as proprietary information. The lack of transparency in the public domain on the calculation of movement "counts" makes it difficult to compare counts as produced by different accelerometer brands or even between versions of the same brand [5,15].

The recent health surveys in the United States and in the UK have used the Actigraph accelerometer as the *de facto standard* device for objective physical activity monitoring [17]. Comparisons between these surveys and future surveys would require either the use of the Actigraph accelerometer or an exact replication of its signal processing method to ensure methodological consistency. Otherwise, observed differences between populations or differences over time may simply be explained by differences in measurement methodology.

Recently, wearable accelerometer devices have become commercially available which are based on MEMS inertial (seismic) acceleration sensors and are capable of storing data at high sampling rates whilst at the same time being feasible for use in epidemiological settings. In this paper, this type of device is referred as a "raw accelerometer" since all frequencies related to human movement are included in the signal without violation of the Nyquist-Shannon sampling theorem. The output data of raw accelerometers is not summarised by any proprietary processing, thus allowing researchers increased control over the data processing stages. In contrast to traditional accelerometers, inertial accelerometers also measure the gravitational component of acceleration and store data in SI units (e.g. $m \cdot s^{-2}$) allowing for between-brand comparisons and easy check of sensor calibration.

The high-resolution raw acceleration data lends itself to a pattern recognition approach for the classification of activity types, which was previously not possible in large-scale studies. This typically involves a feature extraction stage and a prediction stage. Such data processing stages, however, must be carefully described alongside the measurement method for the raw signal since it is the entire inference scheme that determines the validity of the activity classification [21,1].

III. CONVERSION METHOD

The technical specifications of the Actigraph GT1M[1] were used to develop the basics of the method. Mechanical shaker devices have previously been used to test the reproducibility of accelerometers as they allow for testing under controlled motion conditions [2,8,19]. A mechanical shaker was used to empirically assess the response of Actigraph and raw accelerometer devices. This process assisted in the fine-tuning of the conversion method since both devices were subjected to the same type of motion. The proposed conversion method is outlined in Figure 1.

A. Accelerometers

The raw accelerometer device used (GENEA, Unilever Discover, UK) comprises a 3-axis STMicroelectronics accelerometer LIS3LV02DL with a dynamic range of ± 6 g. The acceleration is sampled at 80 Hz and data is locally stored in g unit (1 g = 9.81 m/s^2). The weight is 17 grams (batteries included) and dimensions are 12 x 29 x 37 mm, see Figure 2.

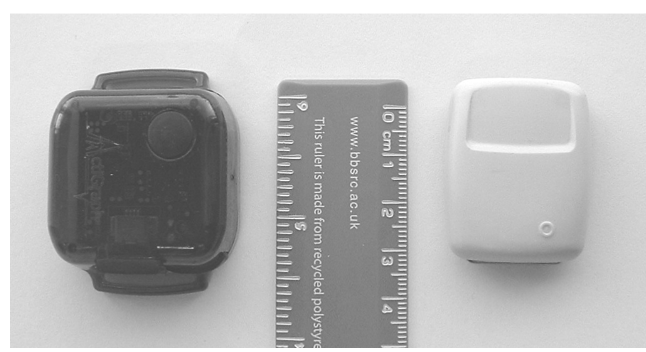

Figure 2. Actigraph and Genea accelerometer

The Actigraph GT1M was used with firmware version 4.3.0. During the mechanical shaker experiments, the GT1M was

[1] http://www.theactigraph.com

978-1-4244-7265-9/10/$26.00 © 2010 IEEE

set in biaxial mode at an epoch length of 1 second; whereas in the free-living activity experiments this device was set in biaxial mode with an epoch length of 5 seconds of which only the vertical counts (AC) were used in the comparison against mimicked counts ($AC_{mimicked}$). These epoch lengths were chosen due to limited memory capacity of the Actigraph GT1M. Fourteen different raw accelerometer devices (GENEA) and twelve Actigraph GT1M were used randomly during the free-living experiments to increase the external validity (generalisability) of the results for each type of accelerometer [8,19].

B. Conversion algorithm development

The technical documentation for the Actigraph GT1M indicates that this device applies a band-pass filter to the acceleration signal with cut-off frequencies set at 0.25 and 2.5 Hz. However, the order of this filter is unspecified. The dynamic range of the sensor is set from 0.05 to 2.5 g according to the technical information. However, the Actigraph contains a dual-axis accelerometer (ADXL320, Analog Devices) with a ±5 g dynamic range. It would seem, therefore, that only one sensor axis is being used and that values are either rescaled or rounded down to 2.5 g in order for the Actigraph GT1M to mimic the signal processing in its predecessor, the Actigraph 7164 [18]. The research literature suggests that counts are likely to be calculated as the area under the filtered and rectified (non-negative) curve [3,4,18]. The gain factor – the ratio between raw acceleration signal (in SI units) and counts - is likely to be brand specific. To understand the various parameters (band-pass filter order and gain factor), a mechanical shaker was used to produce controlled motion conditions (sinusoidal signals) as previously described [2] and shown in Figure 3.

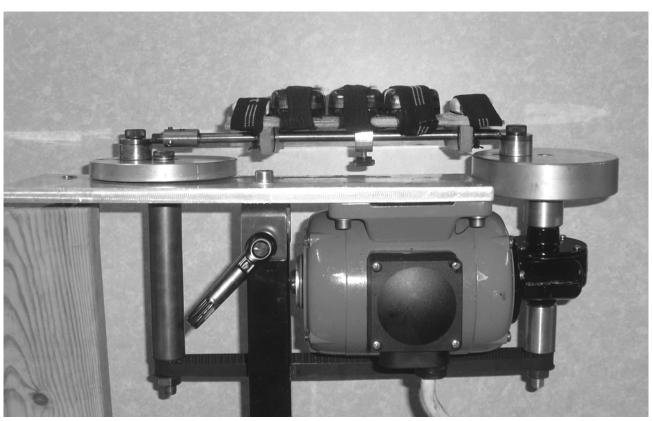

Figure 3. Shaker device

Two raw accelerometer devices (GENEA) and two Actigraph GT1M were attached to the shaker device with their sensors aligned in the plane of shaking, see Figure 3. These four devices were exposed to sixteen frequencies of shaking (0.00, 0.15, 0.20, 0.30, 0.40, 0.56, 0.75, 1.06, 1.31, 1.56, 1.88, 2.13, 2.63, 3.13, 3.63, 4.13 Hz) in a 2-dimensional (horizontal) plane. The frequency distribution was chosen to assess the characteristics of the Actigraph band-pass filter and it was not meant to be representative of the frequency content found in daily physical activities. The experiment was repeated at three radius settings; 22, 35 and 49 mm. Counts were calculated using the processing stages outlined in Figure 1 over two 30-sec intervals within each experimental condition and expressed as counts min^{-1}. The counts derived from the two raw accelerometers, which have two individual axes each oriented parallel to the plane of shaking, were averaged for each experimental condition. The same procedure was applied to the two individual axes of the Actigraph devices. Band-pass filter order was evaluated up to n=4 and that filter order yielding the highest explained variation between Actigraph counts and GENEA raw accelerometer counts was chosen for further analysis. The gain factor between average raw accelerometer counts (un-scaled estimated counts) and Actigraph counts over all conditions was derived.

IV. METHOD EVALUATION

Free-living experiments were conducted to collect accelerometer data for evaluation of the conversion method. Nineteen healthy adult volunteers took part in this study segment, characteristics of whom are described in Table I. Participants were informed about the objective and activity monitoring protocol to be used, in both written and oral forms. All participants gave written informed consent. The Research Ethics Committee of Cambridgeshire approved the study.

	All (N = 19)	**Males** (n =7)	**Females** (N = 12)
Age (y)	47 ±11	49 ±10 (32–64)	43 ±12 (28-58)
Height (cm)	172 ±7	176 ±6 (166-187)	165 ±4 (161-172)
Weight (kg)	73.8 ±13.4	80.3 ±11.9 (67.4-106)	62.8 ±7.7 (52.4-72.5)
BMI ($kg.m^{-2}$)	25.2 ±4.1	26.1 ±4.3 (19.9-35.2)	23.4 ±3.1 (18.6-27.7)

TABLE I. PARTICIPANT CHARACTERISTICS

A. Free-living experiment

Participants were asked to wear one Actigraph on their dominant hip and one GENEA raw accelerometer on their non-dominant hip for seven days in their daily life. The participant was asked to wear the devices during the day but not during bedtime. Raw accelerometer data was converted into Actigraph counts using the processing method presented in Figure 1, and the band-pass filter order and gain factor derived from the mechanical shaker experimental data. The mimicked Actigraph counts ($AC_{mimicked}$) were calculated for 5-second epochs as for Actigraph data but the time-series were compared on a minute-by-minute basis for increased time synchronisation. In addition, the average counts over a week were calculated for each individual.

B. Evaluation Metrics

Explained variation in AC by $AC_{mimicked}$ was assessed through the calculation of the Pearson's correlation coefficient. Reproducibility was assessed using Lin's concordance correlation coefficient, which is the degree to which two readings fall on the 45º line through the origin [11]. To quantify the error of disagreement, the root mean square of the error (RMSE) was computed. Bland & Altman plots were used to assess absolute agreement between the weekly average $AC_{mimicked}$ and AC. A paired sample t-test was applied to test for a significant difference between weekly averages. Linear regression analysis was then used to assess whether the slope and y-intercept of the regression line through average weekly $AC_{mimicked}$ and average weekly AC in daily life differed significantly from respectively one and zero. It is common practice to transform accelerometer time-series data to the frequency domain, so as to describe intensity distribution. Therefore, we also assessed the degree to which the conversion method classifies intensity levels in daily life similarly to the Actigraph for every minute. This is shown graphically by plotting the two intensity distributions (20 bins) and also using the common approach of counting time spent in 4 intensity intervals, defined as: sedentary (0 – 100 counts · min^{-1}), light (100 – 2000 counts · min^{-1}), moderate (2000 – 4000 counts · min^{-1}), and vigorous (> 4000 counts · min^{-1}) intensity. This allows quantifying the sensitivity (degree to which data in an intensity category is classified as such by the mimicking method) and specificity (degree to which data outside an intensity category is classified as such by the mimicking method). All analyses were performed in the open-source tool R using 'signal', 'Hmisc' and 'stat' packages[2]. A significance of $p < .05$ was regarded as acceptable.

V. RESULTS

From the shaker experiment results, a 3rd order Butterworth filter resulted in the highest correlation between Actigraph counts and un-scaled raw accelerometer counts ($r = 0.975$, $p < .01$), followed by a 4th order filter ($r = 0.964$, $p < .01$), and a 2nd order filter ($r = 0.883$, $p < .01$). The gain factor to convert un-scaled raw accelerometer counts into equivalent Actigraph counts was 6.8 (for the frequency range 0.5 – 1.5 Hz which reflect the dominant frequencies in daily human movement).

Average weekly $AC_{mimicked}$ explained 94.2% of the variation in weekly average AC ($p < 0.001$), and for minute-by-minute data, this was 92% ± 5.8% (mean ± SD), ranging from 73.2% to 98.8% ($p < 0.001$). The concordance correlation coefficient for reproducibility was 0.93 (95% C.I.: 0.85 – 0.97) for weekly averages and 0.94 (95% C.I.: 0.94 – 0.94) for minute-by-minute values. Average weekly $AC_{mimicked}$ overestimated AC by 20.9 counts · min^{-1} ($p < .001$). The 95% limits of agreement (representing the range in which 95% of the differences lie) were -15.5 to 57.2 counts · min^{-1}, see Figure 4. Further the RMSE for the weekly average was 27.6 counts · min^{-1} which amounts to 8.39 % when expressed relative to the average.

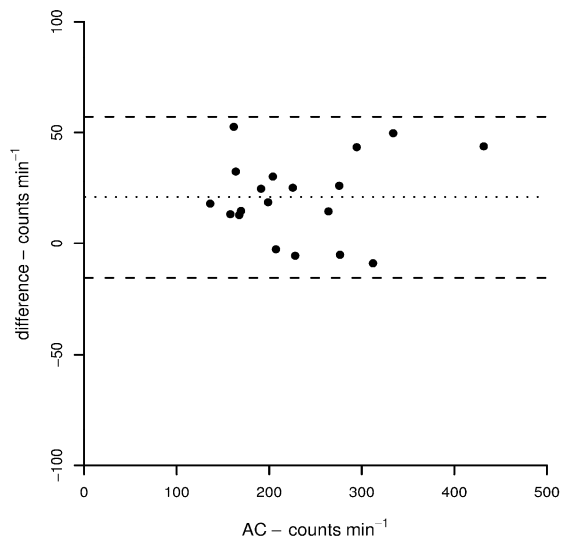

Figure 4. Agreement between weekly average of $AC_{mimicked}$ and AC [lines represent limits of agreement and the mean difference]

The y-intercept of the linear regression line through the average weekly values was 17 and did not significantly differ from zero (95% C.I.: -9 to 43 counts · min^{-1}). The slope of the regression line was 1.01 and did not significantly differ from 1.00 (95% C.I.: 0.91 – 1.13).

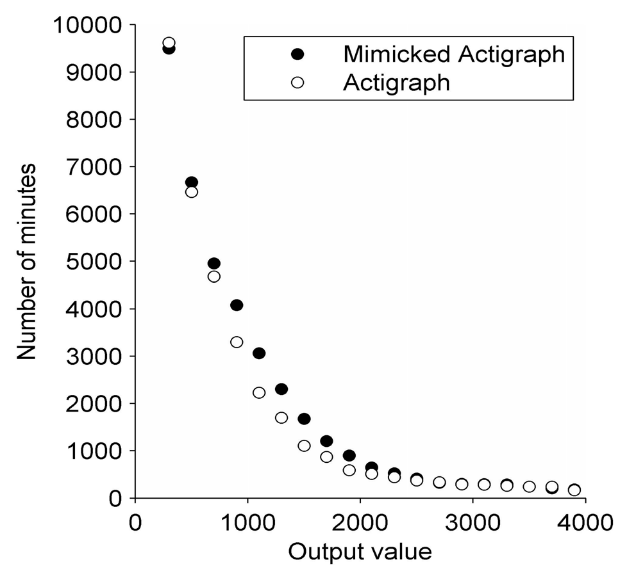

Figure 5. Distribution of output values for Actigraph counts and mimicked Actigraph counts in free-living experiment, excluding the 0-200 interval

[2] http://www.r-project.org/

The distribution of the measurement values for the free-living data showed that the mimicking algorithm assigned more minutes to the range 1000-2000 compared to the Actigraph, see Figure 5. The interval 0-200 was not included in Figure 5 to allow for better visual inspection of the distribution. The number of minutes assigned to the 0-200 interval was 149,569 and 154,652 for $AC_{mimicked}$ and AC respectively.
Sensitivity to discriminate between four intensity levels was > 70.7% and specificity was > 92.5%. These results are presented in Table II.

Level	Sensitivity (%)	Specificity (%)
0 –100 counts min-1	92.7	95.1
100 – 2000 counts min-1	93.4	92.5
2000 – 4000 counts min-1	70.7	99.6
> 4000 counts min-1	81.9	99.9

TABLE II. SENSITIVITY AND SPECIFITY TO ASSESS FOUR INTENSITY LEVELS IN DAILY LIFE (BASED ON 1-MINUTE EPOCHS)

VI. POSTHOC DIAGNOSTIC OF DISAGREEMENT

On the basis of our results, we attempted to diagnose the disagreement. The proposed conversion algorithm assumes a constant gain factor. Figure 6 displays the ratio between un-scaled raw accelerometer counts and Actigraph counts in the shaker experiment. This figure indicates that the gain factor may in fact vary with frequency. However, this pattern may equally well be a consequence of non-perfect matching of the filtering algorithm.

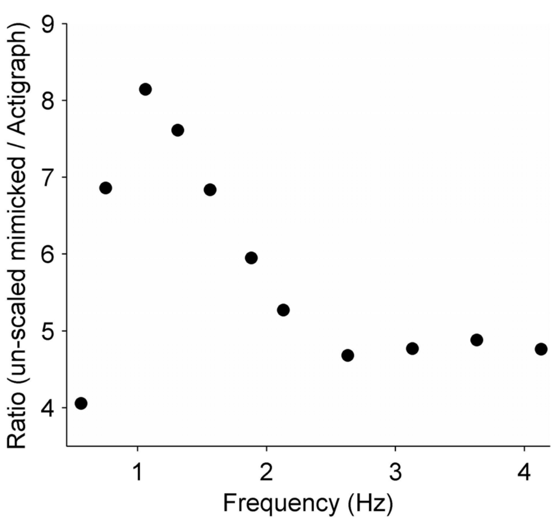

Figure 6. Ratio between un-scaled mimicked counts and Actigraph counts as function of the frequency of shaking

VII. CONCLUSIONS

Actigraph has become the *de facto standard* device for objective physical activity monitoring, and has been used in many large-scale field studies. In this paper, a method to convert raw acceleration data (in S.I. units) into Actigraph counts was proposed and evaluated. Although mimicked counts explained 94.2% of the variation in Actigraph counts, there was a small positive bias and a perfect conversion was not achieved.
Important parameter values were empirically estimated (band-pass filter order and gain factor) and the area under the rectified curve calculation was performed based on the assumption and method described in the literature [18,4]. Although the processing blocks used in the conversion method seem plausible, variations in their characteristics and configuration may be present in the Actigraph. Some of the processing may be performed in the analogue domain which can affect the digital output. It is difficult to perfectly mimic the analogue signal processing by digital processing.
The evaluation of the conversion method in daily physical activity was challenged by the difficulty to isolate differences between the true and mimicked Actigraph counts from: (i) unknown differences in movement between left and right hips and; (ii) unknown differences in the orientation of both accelerometers with regard to the direction of gravity and with regard to the axis of body orientation. The significant overestimation by the raw accelerometer as attached to the non-dominant hip makes it less likely that differences in movement have dominated the difference between the accelerometers.
The conversion method as described in this paper could be seen as the initial step towards the estimation of Actigraph counts from raw accelerometry. With the development of improved techniques raw accelerometer data can be re-analysed to improve compatibility between different types of devices' output. The ability to reanalyze the original raw data demonstrates the flexibility of raw accelerometry compared to the irreversible on-board calculation of counts in traditional accelerometer devices.
Accelerometers have been widely used for the estimation of energy expenditure [20,22,6,7], which is considered a distal outcome. The results of the current study should not be interpreted as the validity of raw accelerometry to estimate energy expenditure or to classify activity types. The validity examined here is that of a more proximal outcome and is subject to characteristics of the acceleration sensors, analogue data processing, on-board digital data processing (if any) and post-hoc digital data processing. Improvements to this validity can be achieved by improvements to each of these components, of which only the latter can be improved after the data has been collected. The current study was carried out using a specific brand of raw accelerometer device but the conversion algorithm can also be applied to other brands based on raw accelerometry (e.g., MTx, Xsens Technologies B.V.). The only requirement here is that the recorded acceleration is expressed in SI units and the dynamic range of the sensor is equal to or larger than ± 2.5 g since this setting was applied in the Actigraph device.

Raw accelerometry can be used to estimate Actigraph counts reasonably well but a perfect replication was not achieved. Industry and research communities could each play important roles in bringing the methodology of physical activity assessment to a higher level, where measurements and derived estimates are more comparable. This would facilitate insights into the importance of physical activity in preserving health at the population level.

VIII. ACKNOWLEDGMENTS

We are grateful to the volunteers for giving their time, the field team of the MRC Epidemiology Unit for assistance in data collection, and to Unilever Discover UK for financial support. We also thanks Andy Hopper for his support.

REFERENCES

[1] Bonomi, A. G., G. Plasqui, et al. "Improving the assessment of daily energy expenditure by identifying types of physical activity using a single accelerometer." J Appl Physiol., 2009

[2] Brage, S., N. Brage, et al. "Reliability and Validity of the Computer Science and Applications Accelerometer in a Mechanical Setting." Measurement in Physical Education and Exercise Science 7(2): 101-119, 2003.

[3] Chang, K. h., M. Y. Chen, et al. "Tracking Free-Weight Exercises", In Proc of UbiComp 2007, Innsbruck, Austria, Berlin, 2007.

[4] Chen, K. Y. and D. R. Bassett, Jr. "The technology of accelerometry-based activity monitors: current and future", Med Sci Sports Exerc 37(11 Suppl): S490-500, 2005.

[5] Corder, K., S. Brage, et al. "Accelerometers and pedometers: methodology and clinical application." Curr Opin Clin Nutr Metab Care 10(5): 597-603, 2007.

[6] Corder, K., U. Ekelund, et al. "Assessment of physical activity in youth." J Appl Physiol 105(3): 977-87, 2008.

[7] Edwards P and Roberts I., ”Population adiposity and climate change”, Int J Epidemiol., 38(4) pp. 1137-40, Aug 2009.

[8] Esliger, D. W. and M. S. Tremblay. "Technical reliability assessment of three accelerometer models in a mechanical setup." Med Sci Sports Exerc 38(12): 2173-81, 2006.

[9] Foresight, "Tackling Obesities: Future Choices", Project Report. Government Office for Science, 2007.

[10] Klesges, R. C., L. M. Klesges, et al. "A validation of two motion sensors in the prediction of child and adult physical activity levels." Am J Epidemiol 122(3): 400-10, 1985.

[11] Lin, L. I. "A concordance correlation coefficient to evaluate reproducibility." Biometrics 45(1): 255-68, 1989.

[12] Lindstrm J et al, ”The Finnish Diabetes Prevention Study (DPS)”, Diabetes Care, 26(12) pp. 3230-6, Dec 2003.

[13] Matthews, C. E. and P. S. Freedson. "Field trial of a three-dimensional activity monitor: comparison with self report." Med Sci Sports Exerc 27(7): 1071-8, 1995.

[14] NHS Information Centre, "Health Survey for England 2005 Latest Trends. NHS Information Centre, 2006.

[15] Rothney, M. P., G. A. Apker, et al. "Comparing the performance of three generations of ActiGraph accelerometers." J Appl Physiol 105(4): 1091-7, 2008.

[16] Rowlands, A. V., M. R. Stone, et al. "Influence of speed and step frequency during walking and running on motion sensor output." Med Sci Sports Exerc 39(4): 716-27, 2007.

[17] Troiano, R. P., D. Berrigan, et al. "Physical activity in the United States measured by accelerometer." Med Sci Sports Exerc 40(1): 181-8, 2008

[18] Tryon, W. W. and R. Williams. "Fully Proportional Actigraphy: A New Instrument." Behavior Research Methods Instruments and Computers 28(3): 392-403, 1996.

[19] Van Hees, V. T., S. M. Slootmaker, et al. "Reproducibility of a triaxial seismic accelerometer (DynaPort)." Med Sci Sports Exerc 41(4): 810-7, 2009.

[20] van Hees, V. T., R. C. van Lummel, et al. "Estimating Activity-related Energy Expenditure Under Sedentary Conditions Using a Tri-axial Seismic Accelerometer." Obesity (Silver Spring), 2009.

[21] Veltink, P. H., H. B. Bussmann, et al. "Detection of static and dynamic activities using uniaxial accelerometers." IEEE Trans Rehabil Eng 4(4): 375-85, 1996.

[22] Westerterp, K. R. "Physical activity assessment with accelerometers." Int J Obes Relat Metab Disord 23 Suppl 3: S45-9, 2009

[23] De Vries SI, Van Hirtum HW, et al. "Validity and reproducibility of motion sensors in youth: a systematic update." Med Sci Sports Exerc. 2009 Apr;41(4):818-27. Review.

[24] Plasqui G, Westerterp KR. "Physical activity assessment with accelerometers: an evaluation against doubly labeled water." Obesity (Silver Spring). 2007 Oct;15(10):2371-9. Review.

[25] Esliger DW, Tremblay MS. "Physical activity and inactivity profiling: the next generation." Can J Public Health. 2007;98 Suppl 2:S195-207. Review.

[26] Preece SJ, Goulermas JY, et al. "Activity identification using body-mounted sensors--a review of classification techniques." Physiol Meas. 30(4): R1-33. 2009. Review.