Prediction of activity related energy expenditure using accelerometer derived physical activity under free-living conditions-a systematic review

Jeran, S., Steinbrecher, A., Pischon, T.

This is the final version of the manuscript. The original article has been published in final edited form in:

International Journal of Obesity
2016 MMM DD ; 40(8): 1187-1197
Advance online publication on 24 May 2016
doi: 10.1038/ijo.2016.14

Publisher: Nature Publishing Group

Prediction of activity related energy expenditure using accelerometer derived physical activity under free-living conditions – a systematic review

Stephanie Jeran[1], Astrid Steinbrecher[1], Tobias Pischon[1,2,3]

[1] Molecular Epidemiology Research Group, Max Delbrück Center for Molecular Medicine (MDC) in the Helmholtz Association, Berlin, Germany

[2] Charité – Universitätsmedizin Berlin, Germany

[3] DZHK (German Center for Cardiovascular Research), partner site Berlin

Keywords: Systematic review, Epidemiology, Human studies, Activity related energy expenditure, Physical activity, Accelerometry

Running title: Prediction of activity related energy expenditure

Corresponding author: Tobias Pischon, Molecular Epidemiology Research Group, Max Delbrück Center for Molecular Medicine (MDC) in the Helmholtz Association, Robert-Rössle-Straße 10, 13125 Berlin, Germany, E-mail: tobias.pischon@mdc-berlin.de, Phone: ++49 30/9406-4563, Fax: ++49 30/9406-4576

Conflict of interest: The authors declare no conflict of interest.

Abstract

Background: Activity related energy expenditure (AEE) might be an important factor in the etiology of chronic diseases. However, measurement of free-living AEE is usually not feasible in large scale epidemiological studies but instead has traditionally been estimated based on self-reported physical activity. Recently, accelerometry has been proposed for objective assessment of physical activity, but it is unclear to what extent this methods explains the variance in AEE.

Methods: We conducted a systematic review searching MEDLINE database (until 2014) on studies that estimated AEE based on accelerometry-assessed physical activity in adults under free-living conditions (using doubly-labeled water method). Extracted study characteristics: sample size, accelerometer (type [uniaxial, triaxial], metrics [e.g. activity counts, steps, acceleration], recording period, body position, wear time), explained variance of AEE (R^2), number of additional predictors. The relation of univariate and multivariate R^2 with study characteristics was analyzed using non-parametric tests.

Results: Nineteen articles were identified. Examination of various accelerometers or subpopulations in one article was treated separately, resulting in 28 studies. Sample sizes ranged from 10-149. In most studies the accelerometer was triaxial, worn at the trunk, during waking hours, and reported activity counts as output metric. Recording periods ranged from 5-15 days. The variance of AEE explained by accelerometer assessed physical activity ranged from 4-80% (median crude R^2 = 26%). Sample size was inversely related to the explained variance. Inclusion of 1 to 3 other predictors in addition to accelerometer output significantly increased the explained variance to a range of 12.5-86% (median total R^2 = 41%). The increase did not depend on the number of added predictors.

Conclusion: We conclude that there is large heterogeneity across studies in the explained variance of AEE when estimated based on accelerometry. Thus, data on predicted AEE based on accelerometry assessed physical activity need to be interpreted cautiously.

Word count: 297

Abbreviations

Acc accelerometer; AEE activity related energy expenditure; BMI body mass index; BMR basal metabolic rate; COPD chronic obstructive pulmonary disease; DIT diet-induced thermogenesis; DLW doubly labeled water; EE energy expenditure; FFM fat-free mass; FM fat mass; IC indirect calorimetry; RMR resting metabolic rate; SD standard deviation; SMR sleeping metabolic rate; TEE total energy expenditure; VM vector magnitude

Introduction

Physical activity is an important factor in the etiology of chronic diseases,[1] whereas less is known about the association between activity related energy expenditure (AEE) and chronic diseases.[2] AEE can be defined as the component of total energy expenditure that is caused by any kind of body movement produced by skeletal muscles.[3] Although AEE is primarily determined by physical activity it also depends on other individual characteristics, such as sex, age, height, or body composition.[4-8] The measurement of AEE under free-living conditions is a challenging task, as AEE is derived as the difference between total energy expenditure (TEE), resting metabolic rate (RMR) (i.e. energy necessary to uphold the basal metabolic functions), and diet induced thermogenesis.[9] The gold standard for measuring TEE under free-living conditions is the doubly labeled water method (DLW), and RMR can be obtained by indirect calorimetry where the amount of oxygen consumption and carbon dioxide production of an individual under fasting and resting conditions is measured.[10]

Nevertheless, these methods are too time and cost intensive to be used in large scale epidemiological studies. Therefore, AEE has been estimated traditionally by relying on questionnaire-based physical activity information, which is then linked to metabolic equivalent (MET) intensity levels to derive an individual’s energy expenditure.[11]

Recently, accelerometry has been introduced into the field of physical activity measurement, and several devices are currently in use in epidemiological studies, such as the German National Cohort.[12] Accelerometers are small devices, which may be attached to the human body for several days and thereby capture objectively acceleration of body movement in up to three planes; thus they can provide information about frequency, intensity and duration of physical activity.[13] Similar to the MET values, the accelerometer output might be used to estimate an individual’s AEE. It is, however, unclear to what extent these devices may explain the variance in AEE under free-living conditions.

Therefore, the aim of this systematic review was to summarize studies that predict AEE based on accelerometry-assessed physical activity data in adults under free-living conditions, and to examine to what extent these predictions explain the variance in AEE. The second aim was to examine to what extent other factors influence these predictions, such as study design, accelerometer device properties or individual characteristics (e.g. age, sex, body composition), and to what extent the addition of such factors improve prediction models.

Methods

Search strategy and study selection

A comprehensive literature search was performed in the MEDLINE database (Medical Literature Analysis and Retrieval System Online) from inception until 2014/12/31 using the following keywords and operators: *energy expenditure* AND (*prediction* OR *estimation* OR *validation* OR *regression* OR *model*) AND (*accelerometry* OR *accelerometer* OR *motion sensor* OR *activity monitor*) AND (*activity* OR *exercise*). The following filters were set: Species: humans; Ages: adult 19+ years; Languages: English, German; Text availability: full text; Publication date: to 2014/12/31. Additionally, reference lists of included articles, and references of reviews and meta-analyses on this topic were hand-searched for further eligible articles.

During the first selection step based on title and abstract, articles were excluded if 1) AEE was not examined, 2) no accelerometer was used to measure physical activity, 3) no adult population aged at least 18 years was examined. All further selection steps were based on full text screening (if available). Articles were excluded for the following reasons: 1) examining no AEE but only TEE, 2) using no accelerometers but pedometers or other devices instead (e.g. heart rate monitors), 3) examining only children or adolescents, 4) setting up the study under no free-living conditions (which requires the use of DLW for measuring TEE), 5) reporting no original data (review articles). One person conducted the literature search and the initial screening of title and abstract, and two people performed the full text screening. Any disagreements were solved by consensus.

Several of the studies that we identified did not aim at prediction of AEE based on accelerometry derived physical activity data but instead compared the AEE estimate that were automatically calculated by the accelerometers based on underlying (in most instances not freely available) algorithms with those AEE derived from DLW. Since the focus of our review was to evaluate the prediction of AEE, we excluded these studies from our analysis. A summary of these studies can be found in Supplementary Table S1.

We considered DLW as the gold standard to assess TEE under free living conditions. The mean difference and its standard deviation of DLW to determine TEE as compared to indirect calorimetry or controlled food intake are reported to range between -2 and 6% and between 1 and 8%, respectively.[14]

Data extraction

From each included article the following information was extracted by two reviewers: first author's name, year of publication; study population characteristics (type of population, sample size, sex, age range or mean, BMI range or mean [or alternatively weight]); accelerometer characteristics (device name, accelerometer type, body position, recording period, wear time); energy expenditure measurement features (period of DLW measurement, measurement or calculation of resting/basal/sleeping metabolic rate [RMR/BMR/SMR], diet induced thermogenesis [DIT], AEE calculation); measured accelerometer output metric; results of association and prediction (if reported): crude explained variance (crude R^2) i.e. the variance in AEE explained solely by accelerometer output; type of prediction model; predictors of final model(s); total R^2 i.e. the variance in AEE explained by accelerometer output and additional predictors; partial R^2 i.e. variance in AEE explained by accelerometer output if other predictors are included in the model; other factors not included in the model. The reported correlation coefficient R was transformed to R^2 if necessary.

If multiple accelerometer devices were used in one article [15-18] or if various subpopulations [19] were examined, each device or population group was regarded as a separate study in our analysis. Among studies that included additional predictors to the association of accelerometer output and AEE, some reported several prediction models.[19-24] From these studies we used only the model that explained the largest variance (total R^2). In cases where the same explained variance was reported for different models we used the model with the lowest number of added predictors (most parsimonious model). In sex-stratified analyses sex was considered as an additional factor,[25] also interactions terms were considered as an additional factor.[19] If mean age was not reported, it was calculated as mean of minimum and maximum of the population's age.[17, 19, 26] Some studies reported various accelerometer output metrics from the same device. Among the different outputs of the accelerometers, the primary output metric that we considered was *counts/day* [16, 18, 23] or *vector magnitude counts/day*.[18, 24] We did not consider *time per intensity categories* as accelerometer outputs. Some studies did not use absolute AEE but (additionally) AEE relative to body weight (AEE/kg).[16, 20, 21, 24, 27-32] If both, absolute und relative AEE, were reported, we considered the absolute value for result description and relation analyses.

Statistical analysis

To assess if study or accelerometer device characteristics influence the association between accelerometer derived physical activity output and DLW derived AEE, we examined the relation between explained variance in AEE and study characteristics in the univariate (crude R^2) and multivariate models (total R^2) graphically and with non-parametric tests. Scatter plots and Spearman rank correlation were used for continuous characteristics (i.e. sample size, mean age), while boxplots and Mann-Whitney *U* test or Kruskal Wallis test were used for categorical characteristics (i.e. accelerometer body position, recording period, wear time, accelerometer output type, accelerometer output metric).

For the subset of studies that reported on adding predictors to the model to improve the explained variance in AEE, we depicted crude and total R^2 with boxplots. We tested improvement of R^2 with Wilcoxon signed rank sum test, and examined the relation between number of additional predictors and total R^2 and improvement of R^2 with Kruskal Wallis test.

In sensitivity analyses we compared the crude explained variance R^2 between the group of studies that reported absolute AEE value and the group that reported AEE relative to body weight (AEE/kg) using the Mann-Whitney *U* test. Studies that reported both values were considered in both groups. We also compared crude R^2 within the group of studies that reported both values (AEE and AEE/kg) using the Wilcoxon signed rank sum test.

All analyses were performed using SAS statistical software, version 9.3 (SAS Institute Inc.). Figures (boxplots, scatter plot) were made using Microsoft Excel 2010 (Microsoft Corporation). Presented P-values are 2-tailed and were considered statistically significant if $p<0.05$.

Results

Search results and study selection

Out of 299 articles resulting from MEDLINE database search, 64 articles remained after reading title and abstracts (Figure 1). After reading full text (not available for 2 articles) 43 articles were excluded. This resulted in 21 articles meeting the inclusion criteria. In addition, 16 articles were included after hand-searching of reference lists of included articles or of reviews or meta-analyses on this topic. Of these 37 articles, 18 were excluded because they did not aim at prediction of AEE but instead on the comparison between accelerometry derived AEE and DLW derived AEE (listed in Supplementary Table S1). Thus, our analysis includes 19 articles, which reported on the prediction of DLW derived AEE based on accelerometry derived physical activity data.

Characteristics of included articles

The characteristics of the 19 included articles are provided in Table 1. Due to the usage of multiple accelerometers in one population [15-18] or the examination of various subpopulations [19] we extracted 28 separate studies on the associations between accelerometer derived physical activity output and measured AEE under free-living conditions. All studies had a cross-sectional study design.

Of these 28 studies, 10 included additional factors beyond accelerometry into the prediction model. Table 2 gives a summary of study characteristics of the 28 studies. Most of the studies were conducted in the general population with sample sizes ranging from 10 to 149 persons. A total of 19 different accelerometer devices were used (7 uniaxial accelerometers, 11 triaxial and 1 biaxial) from a total of 15 different manufacturers. Of the 28 studies, 12 studies applied uniaxial accelerometers, 15 studies applied triaxial accelerometers and 1 study a biaxial accelerometer. Recording periods ranged from 5 to 15 days (Table 2). In most of the studies the accelerometer was worn at the trunk (i.e. hip, lower back, waist or chest; n=20 studies), and the wear time was limited to waking hours (n=21 studies). The most frequently reported accelerometer output metric was uniaxial or triaxial activity counts per time interval.

The measurement period of TEE using the DLW method ranged from 7 to 14 days, and most studies measured resting, basal or sleeping metabolic rate with indirect calorimetry techniques (ventilated hood, handheld mask, respiration chamber) instead of using estimation formulas (Table 1).

Explained variance (R^2) and its relation to study characteristics

In the studies, linear (Pearson correlation, linear regression) and non-linear approaches (Spearman rank correlation, log-linear regression) were used to calculate the variance in DLW derived AEE explained by accelerometer output. We included 24 studies in the analysis, as 4 studies did not report information about crude R^2 values. Crude R^2 ranged from 0.043 to 0.80 (Table 1) with a median of 0.26 (Figure 2). Crude R^2 did not significantly differ by accelerometer recording period (≤1 week vs. >1 week), body position (trunk vs. limbs), wear time (waking hours vs. 24 hours), accelerometer output type (uniaxial vs. triaxial outputs) or accelerometer output metrics (counts vs. steps vs. other) (all p-values of Mann-Whitney *U* test and Kruskal Wallis test >0.05, Figure 2). There was a significant inverse association between crude R^2 and sample size (r= -0.45, p=0.03, Figure 3). There was no significant correlation between crude R^2 and mean age of participants (r=0.16, p=0.44, Figure 3).

<u>Explained variance (R^2) in studies with additional predictors beyond accelerometer output</u>

Ten studies reported on including additional predictors to the association of accelerometer output and DLW derived AEE. The characteristics of these 10 studies were similar to the characteristics of all 28 studies (Table 2). Information about the prediction models of the 10 studies and their performance are listed in Table 3 and summarized in Table 4. The studies by Pomeroy *et al.*[17] were not taken into account in the summary Tables 2 and 4 and in the relation analysis because this study did not report crude R^2 and total R^2.

The studies included between 1 and 3 predictors in addition to accelerometer output metric. Weight and fat-free mass were the most frequently added predictors (Table 4). The explained variance (total R^2) of the multivariate models that included other predictors in addition to accelerometer output ranged from 0.125 to 0.86 (median 0.41), and partial R^2 for accelerometer output ranged from 0.04 to 0.41 (Table 3). Interestingly, in four studies presenting several models,[21-24] those models that included fat-free mass explained a higher proportion of variance in AEE compared to the models that included weight instead.

When stratified by the number of additional predictors, total R^2 did not differ between studies with 1 (n=4, median total R^2=0.42), 2 (n=4, median total R^2=0.37), or 3 additional predictors (n=2, median total R^2=0.63; p=0.56). When examining total R^2 over sample size, there was inverse correlation, similar as with the crude R^2 relation, but without statistical significance (r= -0.53, p=0.12).

Eight studies provided information on crude R^2 (with accelerometer output but no additional predictors) and about total R^2 (with accelerometer output and additional predictors).[19-21, 23-25, 33] Among the 3 studies that added 1 predictor to accelerometer output in the model, the explained variance increased from 0.21 to 0.31 (p=0.25). Among the 3 studies that added 2 predictors, it increased from 0.08 to 0.33 (p=0.25), and among the 2 studies that added 3 predictors, it increased from 0.37 to 0.63 (p=0.50). The improvement of R^2 did not differ between studies with 1 (n=3, median R^2 increase=0.10), 2 (n=3, median R^2 increase=0.26) or 3 additional predictors (n=2, median R^2 increase=0.26; p=0.16). When considering these 8 studies together, the explained variance increased significantly from 0.16 without additional predictors to 0.37 with the largest number of predictors available in each study (p=0.008, Figure 4).

Sensitivity analysis concerning AEE character

Eleven studies reported absolute AEE only as dependent (outcome) variable,[15, 18, 23, 25, 33] 6 studies reported AEE relative to weight (AEE/kg) only,[20, 24, 27-29, 31] and 7 studies reported both, AEE and AEE/kg.[16, 19, 21, 30, 32] There was no difference in crude R^2 between studies using AEE and studies using AEE/kg when analyzed as between-group comparison (AEE studies: n=18, crude R^2 range 0.043 to 0.49, median 0.23; AEE/kg studies: n=13, crude R^2 range 0.05 to 0.80, median 0.35; p=0.09) or as within-group comparison (n=7, AEE studies: crude R^2 range 0.09 to 0.46, median 0.29; AEE/kg studies: crude R^2 range 0.05 to 0.62, median 0.35; p=0.08). Further, similar to our main analysis, there was no relation between crude R^2 and recording period, body position, wear time or accelerometer output type in these two groups. There was also no correlation between crude R^2 and mean age of participants. There was an inverse association between crude R^2 and sample size in both groups, although the correlations were not statistically significant at the 5%-level (AEE studies: r= -0.37, p=0.13; AEE/kg studies: r= -0.51, p=0.07).

Discussion

In this systematic review, we identified 19 articles resulting in 28 underlying studies that estimated AEE based on accelerometry-assessed physical activity data under free-living conditions in a general adult population. The explained variance of DLW derived AEE from single accelerometer output was quite broad and ranged from 4-80%. Sample size was the only parameter that was related to the explained variance across studies, in a way that it was lower in studies with larger sample size. Parameters such as accelerometer output type or output metrics, recording period, body position of the accelerometer, wear time, or age did not systematically explain this heterogeneity. Inclusion of predictors other than accelerometry significantly improved the explained variance in AEE, although this did not depend on the number of predictors included.

We speculate that the heterogeneity observed in our review for the explained variance in AEE across studies is partly due to different study designs. This may be supported by our observation that studies with smaller sample size resulted in higher explained variances, whereas studies with larger sample size resulted in lower explained variances. Smaller studies often include volunteers selected from a special group (e.g. conscripts[26] or elderly[15, 16]), where the range of personal characteristics that may impact on energy expenditure (such as age, weight, height, body composition or activity patterns) is likely to be smaller than in larger studies. Accordingly, the variance in AEE that is due to these personal characteristics will be smaller, and, as a consequence, the variance explained by accelerometry assessed physical activity will be greater in smaller studies than in larger studies.

Furthermore, accelerometers might also detect different types of activities with different accuracy;[13] e.g. accelerometers worn at the hip may be more likely to detect activities that involve movement of the trunk but less likely to detect activities that involve movements of the arms. We speculate that the type of activities in larger studies might be more heterogeneous as compared with smaller studies. This could result in smaller variance in AEE based on accelerometry in larger studies with more heterogeneous types of activities. Unfortunately, the data we obtained from the published studies included in our systematic review did not allow to investigate the effect of type of activities on the explained variance in more detail.

In addition, it is likely that different procedures of data processing may further contribute to heterogeneity across studies. On the one hand there is a large variety of available accelerometer devices with each manufacturer having its own approach to filter, amplify or convert the acceleration signals into an output value, commonly activity counts, that is, however, not comparable between different accelerometers.[34] Unfortunately, this information on data processing was not readily available from the publications included in our analysis, and, therefore, we were not able to investigate whether such variety may systematically account for the differences in the explained variances across studies. We were also not able to analyze data by manufacturer or device model which might have been a proxy for different data processing techniques,[35, 36] because the multitude of manufacturers and device models did not allow an aggregated analysis.

Further, the length of the recording period per day may influence the variance of AEE explained by accelerometry assessed physical activity. Overall, we found no significant differences between studies that applied accelerometry 24 hours per day and those that recorded physical activity during waking hours only. However, the definition of waking hours was different across studies and may further contribute to heterogeneity of results. For example, Herrmann et al. showed that estimation of daily physical activity based on extrapolation of step counts recorded for less than 12 hours/day may underestimate “true” average daily physical activity.[37] The placement of the accelerometer at the body could determine whether and how valid all the different activities are detected, and the length of recording period should cover the typical activity pattern of a person but without having a negative impact on the compliance.[38] Nevertheless, we found no significant differences in the explained variance of AEE for studies with different body placements or with different recording periods.

We found no significant differences in the explained variance of AEE between studies that used triaxial accelerometer outputs versus those that used uniaxial outputs. Theoretically, triaxial accelerometers may record physical activity with higher validity than uniaxial devices.[6, 38, 39] Further, some studies suggest that activities in sedentary or standing postures may be detected more sufficiently with triaxial as compared to uniaxial accelerometers.[9, 38] In our review, two articles analyzed both the uniaxial and the triaxial output of one used monitor.[18, 24] In both studies the explained variance of AEE was slightly higher for the triaxial output compared to the uniaxial (R^2=0.81 vs. 0.77 [24]; R^2=0.29 vs. 0.27 [18]). Since these differences are rather small we speculate that both accelerometer types seem to provide comparable information about AEE in free-living subjects.[38]

After inclusion of other predictors in addition to accelerometer output into the prediction models the explained variance in AEE generally increased and ranged from 12.5-86%, which is still quite a broad range. We found no clear association between the number of additional included predictors and the explained variance of that model. Likewise, the improvement of the explained variance after including additional predictors did not differ with the number of additional predictors. The partial explained variance for accelerometer output ranged from 4-41%. We therefore speculate that not the number of predictors is most important but which predictors are included. However, the number of studies was low in this analysis, and therefore results have to be interpreted cautiously.

Physiologically plausible predictors are body weight or fat-free mass, as they are associated with energy expenditure and physical activity.[7] In our review, these were the most frequently used predictors. Interestingly, when fat-free mass was included in the model instead of weight the explained variance slightly increased.[21-24] This could be explained by the different impact of fat-free mass and fat mass or weight on AEE and physical activity.[24, 40] This again illustrates the difference between AEE and physical activity: Two persons of same physical activity and same weight may have different amount of AEE due to differences in body composition, because the impact of fat-free mass on AEE as metabolic active component of weight is higher as the impact of fat mass.[7] Other factors, also associated with energy expenditure, like height, age, sex were less often included in the prediction models.

As already mentioned, the comparison of accelerometer outputs between the studies is limited due to the arbitrary character of "counts".[13] Therefore, the comparison of the improvement by included additional predictors should be interpreted cautiously.

The strength of this review is that we focused only on studies under free-living conditions using DLW that examined the association of accelerometer derived physical activity output. Other reviews also included studies that compared accelerometry derived AEE with DLW derived AEE or studies under laboratory conditions with predefined activity protocols,[4-6, 41] that could mislead the interpretation of association and prediction of AEE by accelerometry in free-living populations. For comparison we listed those identified studies from our search in the Supplementary Table S1.

Our review also has some limitations. Overall, the number of studies we identified was relatively small, and there was substantial heterogeneity in study characteristics. For example, most studies included in our review provided information about absolute AEE but we also included the few studies that reported only AEE relative to body weight. In sensitivity analyses, the explained variance of AEE was slightly but not significantly higher for studies on AEE relative to body weight as compared to studies on absolute AEE. This is in line with our speculation that body weight is an important factor to consider when predicting AEE; either by including it as a prediction factor or by standardizing AEE on body weight.

Further, in two articles information of accelerometer type, body position or wear time was not reported,[16, 31] and we therefore made assumptions based on other references using the same device in order to compensate the missing information. Next, in three studies the periods of DLW measurement and accelerometry recording didn't strictly overlap.[22, 30, 33] Overall, these often subtle differences across studies may have increased the heterogeneity observed for the variance of AEE explained by accelerometry that we found in our review. In addition, the pooling of heterogeneous studies into groups may attenuate between-group differences towards the null.

Another methodological limitation is the strict choice of some search terms so that many studies were only identified by hand-searching the reference lists.

In order to improve the prediction of free-living AEE based on accelerometry derived physical activity, in theory there are at least two major possibilities: One is to improve the assessment of physical activity by use of technically more advanced accelerometers;[34] the other one is to consider factors other than physical activity that may affect AEE. Unfortunately, based on our review, current studies that have been published so far do not allow to make evidence-based recommendations on how to improve prediction. Therefore, future studies are urgently needed to investigate in detail how AEE can best be predicted based on accelerometry, and which factors should be considered.

In conclusion, we found a large heterogeneity across studies in the explained variance of models that predict AEE based on accelerometry data assessed in persons under free-living conditions. The explained variance was smaller in studies with larger sample size. Addition of factors other than accelerometry significantly improved the prediction but this improvement did not depend on the number of factors added. These data indicate that AEE as estimated based on accelerometry needs to be interpreted cautiously. Further development of prediction models in population based studies under free-living conditions is needed, with focus on improved and comparable measurement of physical activity by accelerometry.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

All authors designed the study. SJ conducted the search and screened the abstracts; SJ and AS conducted full text screening and data extraction. SJ analyzed the data and all authors interpreted the data. SJ drafted the manuscript. All authors revised and approved the final version.

Supplementary Information

Additional supplementary information is available in the online version of this article on *International Journal of Obesity* website.

Table S1. Characteristics and results of articles on comparison of AEE estimates of accelerometer and DLW derived AEE

References

1. World Health Organization. *Global Recommendations on Physical Activity for Health*. Geneva: World Health Organization, 2010.
2. Zheng C, Beresford SA, Van Horn L, Tinker LF, Thomson CA, Neuhouser ML *et al.* Simultaneous association of total energy consumption and activity-related energy expenditure with risks of cardiovascular disease, cancer, and diabetes among postmenopausal women. *Am J Epidemiol* 2014; **180**: 526-535.
3. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. *Public Health Rep* 1985; **100**: 126-131.
4. Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. *Obesity (Silver Spring)* 2007; **15**: 2371-2379.
5. Plasqui G, Bonomi AG, Westerterp KR. Daily physical activity assessment with accelerometers: new insights and validation studies. *Obes Rev* 2013; **14**: 451-462.
6. Van Remoortel H, Giavedoni S, Raste Y, Burtin C, Louvaris Z, Gimeno-Santos E *et al.* Validity of activity monitors in health and chronic disease: a systematic review. *Int J Behav Nutr Phys Act* 2012; **9**: 84.
7. Westerterp KR. Physical activity and physical activity induced energy expenditure in humans: measurement, determinants, and effects. *Front Physiol* 2013; **4**: 90.
8. Westerterp KR. Physical activity as determinant of daily energy expenditure. *Physiol Behav* 2008; **93**: 1039-1043.
9. Westerterp KR. Assessment of physical activity: a critical appraisal. *Eur J Appl Physiol* 2009; **105**: 823-828.
10. Pinheiro Volp AC, Esteves de Oliveira FC, Duarte Moreira Alves R, Esteves EA, Bressan J. Energy expenditure: components and evaluation methods. *Nutr Hosp* 2011; **26**: 430-440.

11. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Jr., Tudor-Locke C *et al.* 2011 Compendium of Physical Activities: a second update of codes and MET values. *Med Sci Sports Exerc* 2011; **43**: 1575-1581.
12. German National Cohort Consortium. The German National Cohort: aims, study design and organization. *Eur J Epidemiol* 2014; **29**: 371-382.
13. Warren JM, Ekelund U, Besson H, Mezzani A, Geladas N, Vanhees L. Assessment of physical activity - a review of methodologies with reference to epidemiological research: a report of the exercise physiology section of the European Association of Cardiovascular Prevention and Rehabilitation. *Eur J Cardiovasc Prev Rehabil* 2010; **17**: 127-139.
14. Speakman JR. The history and theory of the doubly labeled water technique. *Am J Clin Nutr* 1998; **68**: 932S-938S.
15. Rabinovich RA, Louvaris Z, Raste Y, Langer D, Van Remoortel H, Giavedoni S *et al.* Validity of physical activity monitors during daily life in patients with COPD. *Eur Respir J* 2013; **42**: 1205-1215.
16. Colbert LH, Matthews CE, Havighurst TC, Kim K, Schoeller DA. Comparative validity of physical activity measures in older adults. *Med Sci Sports Exerc* 2011; **43**: 867-876.
17. Pomeroy J, Brage S, Curtis JM, Swan PD, Knowler WC, Franks PW. Between-monitor differences in step counts are related to body size: implications for objective physical activity measurement. *PLoS One* 2011; **6**: e18942.
18. Leenders NY, Sherman WM, Nagaraja HN, Kien CL. Evaluation of methods to assess physical activity in free-living conditions. *Med Sci Sports Exerc* 2001; **33**: 1233-1240.
19. van Hees VT, Renstrom F, Wright A, Gradmark A, Catt M, Chen KY *et al.* Estimation of daily energy expenditure in pregnant and non-pregnant women using a wrist-worn tri-axial accelerometer. *PLoS One* 2011; **6**: e22922.
20. Assah FK, Ekelund U, Brage S, Corder K, Wright A, Mbanya JC *et al.* Predicting physical activity energy expenditure using accelerometry in adults from sub-Sahara Africa. *Obesity (Silver Spring)* 2009; **17**: 1588-1595.
21. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. Estimation of free-living energy expenditure using a novel activity monitor designed to minimize obtrusiveness. *Obesity (Silver Spring)* 2010; **18**: 1845-1851.
22. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *J Appl Physiol (1985)* 2009; **107**: 655-661.
23. Masse LC, Fulton JE, Watson KL, Mahar MT, Meyers MC, Wong WW. Influence of body composition on physical activity validation studies using doubly labeled water. *J Appl Physiol (1985)* 2004; **96**: 1357-1364.
24. Plasqui G, Joosen AM, Kester AD, Goris AH, Westerterp KR. Measuring free-living energy expenditure and physical activity with triaxial accelerometry. *Obes Res* 2005; **13**: 1363-1369.
25. Horner F, Bilzon JL, Rayson M, Blacker S, Richmond V, Carter J *et al.* Development of an accelerometer-based multivariate model to predict free-living energy expenditure in a large military cohort. *J Sports Sci* 2013; **31**: 354-360.
26. Kinnunen H, Tanskanen M, Kyrolainen H, Westerterp KR. Wrist-worn accelerometers in assessment of energy expenditure during intensive training. *Physiol Meas* 2012; **33**: 1841-1854.
27. Valenti G, Camps SG, Verhoef SP, Bonomi AG, Westerterp KR. Validating measures of free-living physical activity in overweight and obese subjects using an accelerometer. *Int J Obes (Lond)* 2014; **38**: 1011-1014.
28. Skipworth RJ, Stene GB, Dahele M, Hendry PO, Small AC, Blum D *et al.* Patient-focused endpoints in advanced cancer: criterion-based validation of accelerometer-based activity monitoring. *Clin Nutr* 2011; **30**: 812-821.
29. Assah FK, Ekelund U, Brage S, Wright A, Mbanya JC, Wareham NJ. Accuracy and validity of a combined heart rate and motion sensor for the measurement of free-living physical activity energy expenditure in adults in Cameroon. *Int J Epidemiol* 2011; **40**: 112-120.

30. Pietilainen KH, Kaprio J, Borg P, Plasqui G, Yki-Jarvinen H, Kujala UM *et al.* Physical inactivity and obesity: a vicious circle. *Obesity (Silver Spring)* 2008; **16**: 409-414.
31. Adams SA, Matthews CE, Ebbeling CB, Moore CG, Cunningham JE, Fulton J *et al.* The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol* 2005; **161**: 389-398.
32. Bouten CV, Verboeket-van de Venne WP, Westerterp KR, Verduin M, Janssen JD. Daily physical activity assessment: comparison between movement registration and doubly labeled water. *J Appl Physiol (1985)* 1996; **81**: 1019-1026.
33. Tudor-Locke C, Martin CK, Brashear MM, Rood JC, Katzmarzyk PT, Johnson WD. Predicting doubly labeled water energy expenditure from ambulatory activity. *Appl Physiol Nutr Metab* 2012; **37**: 1091-1100.
34. Chen KY, Bassett DR, Jr. The technology of accelerometry-based activity monitors: current and future. *Med Sci Sports Exerc* 2005; **37**: S490-500.
35. Esliger DW, Tremblay MS. Technical reliability assessment of three accelerometer models in a mechanical setup. *Med Sci Sports Exerc* 2006; **38**: 2173-2181.
36. Welk GJ, McClain J, Ainsworth BE. Protocols for evaluating equivalency of accelerometry-based activity monitors. *Med Sci Sports Exerc* 2012; **44**: S39-49.
37. Herrmann SD, Barreira TV, Kang M, Ainsworth BE. Impact of accelerometer wear time on physical activity data: a NHANES semisimulation data approach. *Br J Sports Med* 2014; **48**: 278-282.
38. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc* 2005; **37**: S531-543.
39. Vanhelst J, Beghin L, Duhamel A, Bergman P, Sjostrom M, Gottrand F. Comparison of uniaxial and triaxial accelerometry in the assessment of physical activity among adolescents under free-living conditions: the HELENA study. *BMC Med Res Methodol* 2012; **12**: 26.
40. Prentice AM, Goldberg GR, Murgatroyd PR, Cole TJ. Physical activity and obesity: problems in correcting expenditure for body size. *Int J Obes Relat Metab Disord* 1996; **20**: 688-691.
41. Westerterp KR. Reliable assessment of physical activity in disease: an update on activity monitors. *Curr Opin Clin Nutr Metab Care* 2014; **17**: 401-406.

Figure legends

Figure 1 Flow chart of article selection

Abbreviations: AEE activity related energy expenditure; DLW doubly labeled water; EE energy expenditure

Figure 2 Crude R^2 of all 24 studies, and stratified by accelerometer characteristics (recording period, body position, wear time, accelerometer output type, accelerometer output metric); circle = outlier; *) category includes the categories “acceleration” and “other” provided in Table 2

Figure 3 Scatterplots of crude R^2 and sample size or mean age based on 24 studies

Figure 4 Crude and total R^2 for subset of studies adding additional predictors (n=8); circle = outlier

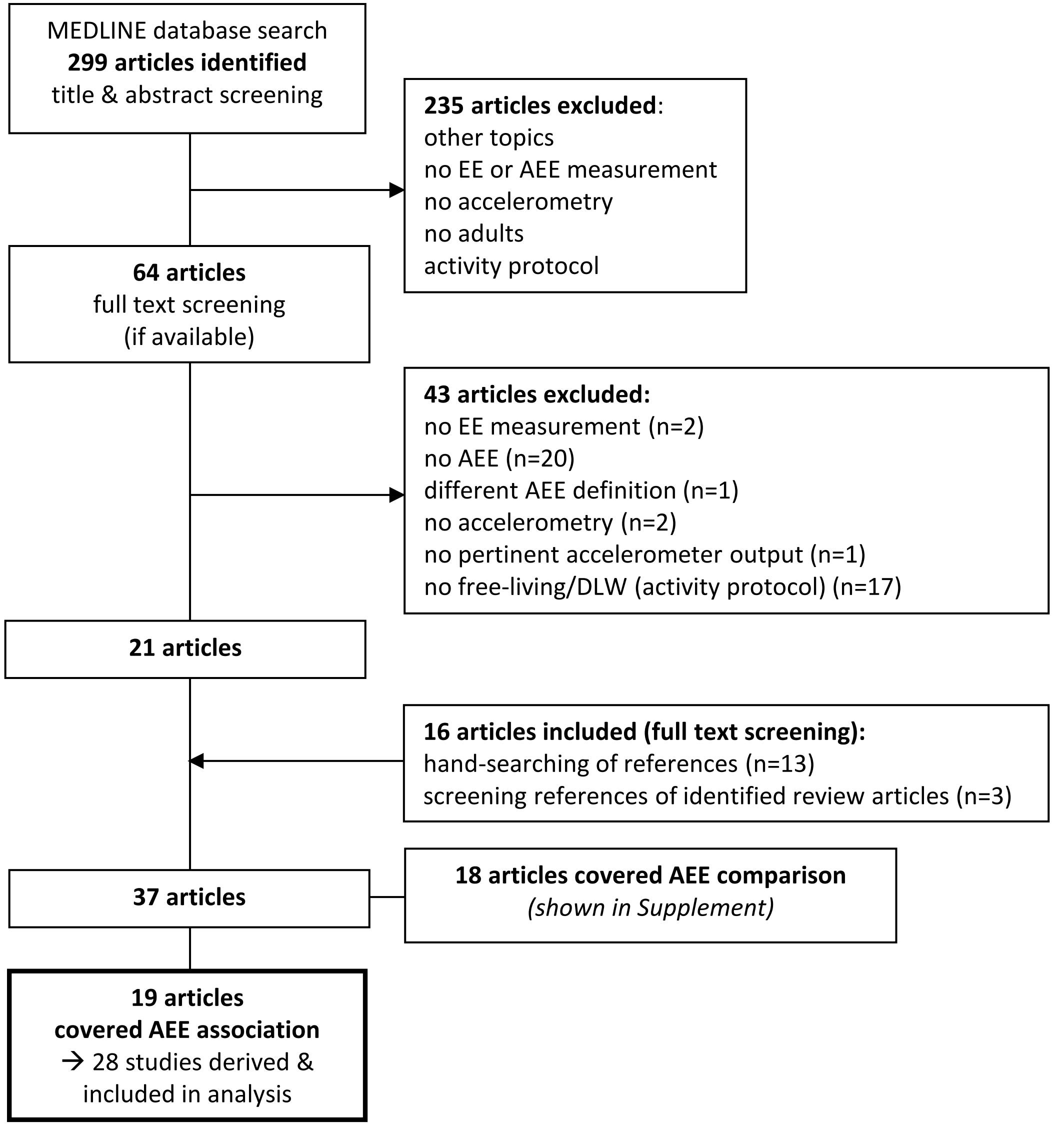
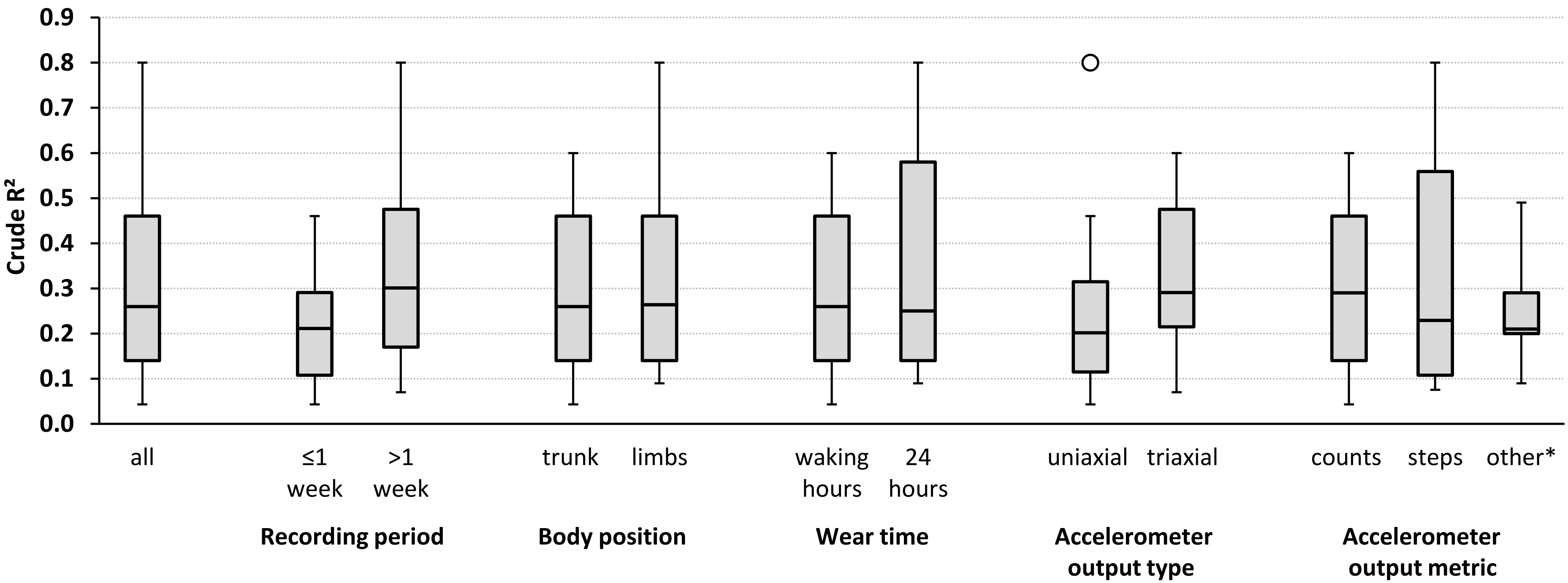
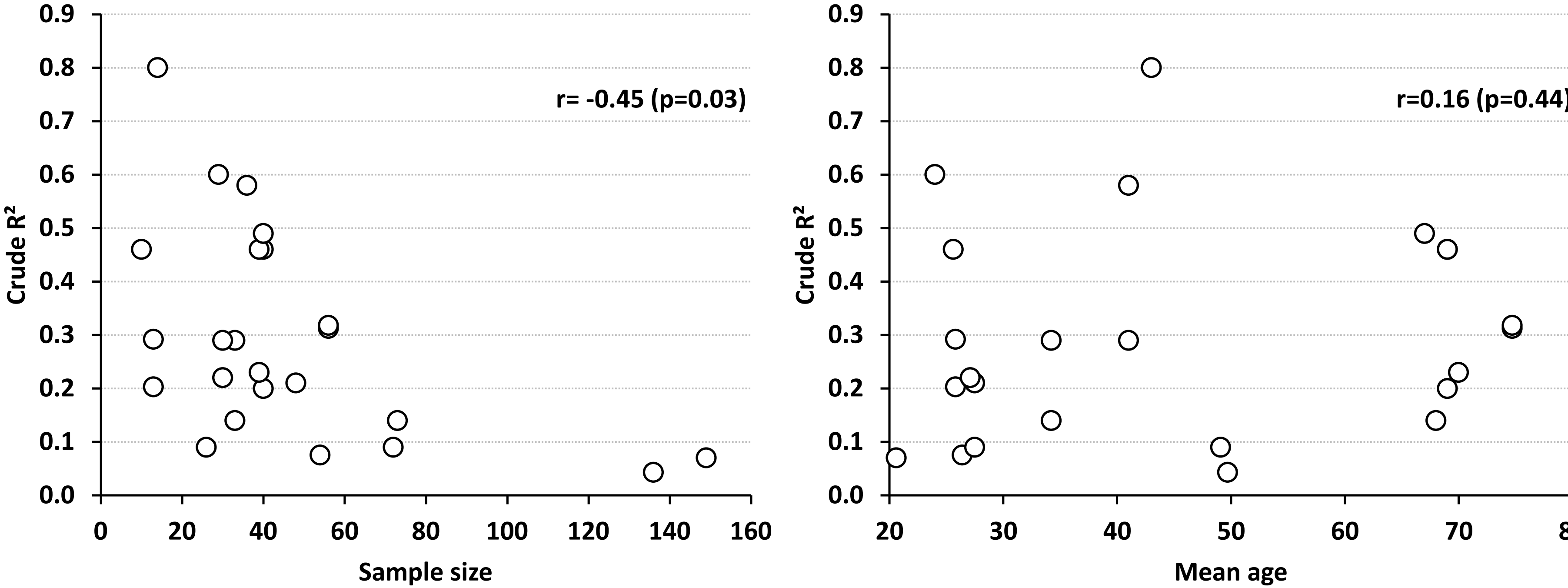

MEDLINE database search
299 articles identified
title & abstract screening
235 articles excluded:
other topics
no EE or AEE measurement
no accelerometry
no adults
activity protocol
64 articles
full text screening
(if available)
43 articles excluded:
no EE measurement (n=2)
no AEE (n=20)
different AEE definition (n=1)
no accelerometry (n=2)
no pertinent accelerometer output (n=1)
no free-living/DLW (activity protocol) (n=17)
21 articles
16 articles included (full text screening):
hand-searching of references (n=13)
screening references of identified review articles (n=3)
37 articles
18 articles covered AEE comparison
(shown in Supplement)
19 articles
covered AEE association
→ 28 studies derived &
included in analysis

figure1



Crude R²
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
all
≤1 week
>1 week
trunk
limbs
waking hours
24 hours
uniaxial
triaxial
counts
steps
other*
Recording period
Body position
Wear time
Accelerometer output type
Accelerometer output metric

Figure2



Crude R²
Sample size
r= -0.45 (p=0.03)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
20
40
60
80
100
120
140
160
Crude R²
Mean age
r=0.16 (p=0.44)
20
30
40
50
60
70
80

figure3

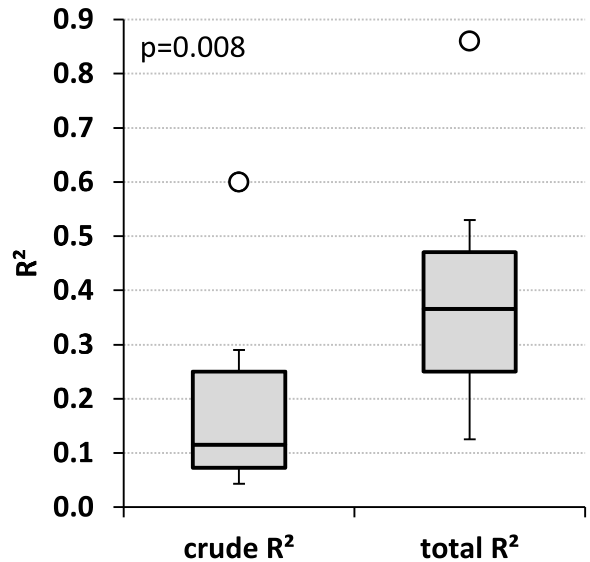


figure4

Table 1 Characteristics and results of 19 included articles

<u>Reference</u>	<u>Study participants</u> **Population** **Sample size (n)** **Sex (male/female)** **Age range** (or mean ±SD) **BMI/weight range** (or mean ±SD)	<u>Accelerometer</u> **Device name** **Acc type** **Body position** **Recording period** **Wear time**	<u>Energy expenditure</u> **DLW measurement period** **RMR/BMR/SMR** measurement or calculation (equation) **DIT included** **AEE calculation**	**Measured Accelerometer output metric**	**Crude R^2** Acc-output vs. AEE (§) AEE/kg (#)
Rabinovich *et al.*, 2013 [15]	COPD patients[a] n=40 28 male, 12 female 69 ±5.8 years 27.2 ±4.7 kg/m²	Lifecorder PLUS uniaxial left waist 14 days, waking hours	TEE: DLW (14 days) RMR: IC (hood) DIT: assumed 10% TEE AEE = TEE*0.9-RMR	Activity score	§ R^2=0.20
	COPD patients[a] n=40 29 male, 11 female 69 ±6.1 years 26.8 ±4.8 kg/m²	Actiwatch Spectrum uniaxial left wrist 14 days, waking hours		Activity counts	§ R^2=0.46
	COPD patients[a] n=39 27 male, 12 female 70 ±5.9 years 25.9 ±5.2 kg/m²	RT3 triaxial right waist 14 days, waking hours		VM units	§ R^2=0.23
	COPD patients[a] n=40 33 male, 7 female 67 ±6.2 years 27.2 ±4.1 kg/m²	DynaPort Move Monitor triaxial waist, lower back 14 days, waking hours		Movement intensity	§ R^2=0.49
	COPD patients[a] n=39 32 male, 7 female 69 ±6.6 years 26.8 ±4.5 kg/m²	ActiGraph GT3X triaxial right waist 14 days, waking hours		VM units	§ R^2=0.46
	COPD patients[a] n=73 58 male, 15 female 68 ±6.4 years 26.4 ±4.5 kg/m²	SenseWear Armband Mini triaxial upper left arm, triceps 14 days, waking hours		Steps	§ R^2=0.14
Valenti *et al.*, 2013 [27]	overweight, obese n=36 11 male, 25 female 41 ±7 years 31.0 ±2.5 kg/m²	TracmorD triaxial lower back 14 days, 24 hours	TEE: DLW (14 days) SMR: IC (chamber) DIT: assumed 10% TEE AEE = 0.9*TEE-SMR	Counts/day	# R^2=0.58
Horner *et al.*, 2013 [25]	10 military cohorts participants n=149 108 male, 41 female 20.6 ±3.9 years <u>weight</u> 67.9 ±12.0 kg	3DNX model V3 & V2 triaxial small of the back 7-10 days, waking hours	TEE: DLW (7-10 days) RMR: calc (Schofield) DIT: assumed 10% TEE AEE = TEE*0.9-RMR	log Counts/day	R^2=0.07 (vs. log AEE)
Tudor-Locke *et al.*, 2012 [33]	normal- & overweight n=54 20 male, 34 female 20-36 years 18.5-27.6 kg/m²	Actigraph GT1M uniaxial hip 7 days, waking hours	TEE: DLW (14 days) RMR: IC (hood) DIT: assumed 10% TEE AEE = TEE*0.9-RMR	Steps/day	§ R^2=0.0752

Kinnunen *et al.*, 2012 [26]	Conscripts n=22 male 19-20 years <u>weight</u> 57-111 kg	Polar activity recorder uniaxial non-dominant wrist 7 days, 24 hours	TEE: DLW (7 days) BMR: calc (Schofield) DIT: assumed 10% TEE AEE = TEE*0.9-BMR	Normalized hand movements/ min	n.s.
Skipworth *et al.*, 2011 [28]	cancer patients, healthy subjects n=14 12 male, 2 female 25-76 years 20.4-33.5 kg/m²	ActivPAL uniaxial right thigh 14 days, 24 hours	TEE: DLW (14 days) RMR: IC (hood) AEE=TEE-RMR	Steps/min	# R^2=0.80
van Hees *et al.*, 2011 [19]	non-pregnant n=65[b] female 20-35 years 27.8 ±6.6 kg/m²	GENEA triaxial right or left wrist 10 days, 24 hours	TEE: DLW (10 days) RMR: IC (hood) DIT: assumed 10% TEE AEE = 0.9*TEE-RMR	VM-acceleration[c]	§ R^2=0.21 # R^2=0.27[d] (n=48, >7 days)
	pregnant n=30[b] female 20-35 years 27.7 ±5.3 kg/m²				§ R^2=0.09 # R^2=0.05[d] (n=26, >7 days)
Colbert *et al.*, 2011 [16]	older adults n=56 12 male, 44 female 74.7 ±6.5 years 25.8 ±4.2 kg/m²	Actigraph GT1M uniaxial right waist 10 days, waking hours	TEE: DLW (10 days) RMR: IC (hood) DIT: assumed 10% TEE AEE = 0.9*TEE-RMR	1) Steps/day 2) Counts/day	1) § R^2=0.342 # R^2=0.396 2) § R^2=0.312 # R^2=0.315
		SenseWear Pro3 Armband n.s. (biaxial)[e] n.s. (upper arm)[e] 10 days, waking hours		Steps/day	§ R^2=0.318 # R^2=0.348
Pomeroy *et al.*, 2011 [17]	American Indians n=50 25 male, 25 female 20-34 years 30.0 (men), 25.6 (women) kg/m² (median)	ActiGraph MTI (model 7164) uniaxial hip 7 days, waking hours	TEE: DLW (7 days) RMR: IC (hood) DIT: assumed 10% TEE AEE = 0.9*TEE-RMR	Steps/day	n.s.
		Dynastream AMP-331 triaxial ankle 7 days, waking hours		Steps/day	n.s.
Assah *et al.*, 2011 [29]	healthy urban/rural Cameroonians n=33 16 male, 17 female 25-50 years 27.1 ±4.6 kg/m²	Actiheart uniaxial + heart rate chest 7 days, 24 hours	TEE: DLW (7 days) RMR: IC (handheld) SMR: calc (=0.9*RMR) DIT assumed 10% TEE AEE=0.9*TEE-RMR-SMR	Acceleration	# R^2=0.29
Bonomi *et al.*, 2010 [21]	healthy volunteers n=30 18 male, 12 female 26-60 years 19.0-31.4 kg/m²	TracmorD triaxial lower back 14 days, waking hours	TEE: DLW (14 days) SMR: IC (chamber) DIT: assumed 10% TEE AEE = TEE*0.9-SMR	Counts/day	§ R^2=0.29 # R^2=0.50
Bonomi *et al.*, 2009 [22]	healthy volunteers n=15 9 male, 6 female 26-59 years 19.6-29.5 kg/m²	Tracmor triaxial lower back 5 days, waking hours	TEE: DLW (14 days) SMR: IC (chamber) DIT: assumed 10% TEE AEE = 0.9*TEE-SMR	Counts/day	n.s.

Assah *et al.*, 2009 [20]	healthy urban/rural Cameroonians n=33 16 male, 17 female 25-50 years 27.1 ±4.6 kg/m²	Actigraph GT1M uniaxial waist 7 days, 24 hours	TEE: DLW (7 days) RMR: IC (handheld) SMR: calc (=0.9*RMR) DIT: assumed 10% TEE AEE = TEE-RMR-SMR-DIT	Counts/day	# R^2=0.14
Pietiläinen *et al.*, 2008 [30]	monozygotic twins (weight discordant) n=20[b] 10 male, 10 female 25.6 ±1.3 years[f] 25.7 ±2.7 kg/m² (non obese), 31.4 ±2.2 kg/m² (obese)[f]	Tracmor triaxial lower back 7 days, waking hours	TEE: DLW (14 days) BMR: IC (hood) DIT: assumed 10% TEE AEE = 0.9*TEE-BMR	Counts/day	§ R^2=0.46 # R^2=0.62 (n=10)
Plasqui *et al.*, 2005 [24]	healthy twins n=29 10 male, 19 female 24 ±6 years 22.9 ±4.3 kg/m²	Tracmor triaxial lower back 15 days, waking hours	TEE: DLW (14 days) SMR: IC (chamber) DIT: assumed 10% TEE AEE = 0.9*TEE-SMR	VM-counts/day (vertical counts/day)	# R^2=0.60 (n.s.)
Adams *et al.*, 2005 [31]	university & general population n=80[b] female 40-65 years 18.7-38.2 kg/m²	Actigraph MTI (model 7164) uniaxial n.s. (trunk)[e] 14 days, n.s. (waking hours)[e]	TEE: DLW (14 days) RMR: calc (Arciero) AEE = TEE-RMR	Daily counts/ min	# R^2=0.09 (n=72)
Mâsse *et al.*, 2004 [23]	healthy African American, Hispanic n=136 female 40.1-71.1 years 19.1-59.9 kg/m²	CSA (model 7164) uniaxial right hip 7 days, waking hours	TEE: DLW (14 days, second week used) RMR: IC (hood) DIT: assumed 10% TEE AEE = TEE-DIT-RMR	1) Counts/day 2) Weight adj. counts/day 3) FFM adj. counts/day	1) § R^2=0.043 2) § R^2=0.090 3) § R^2=0.109
Leenders *et al.*, 2001 [18]	healthy volunteers n=13 female 21-37 years 19.9-27.7 kg/m²	Tritrac-R3D triaxial right hip 7 days, waking hours	TEE: DLW (7 days) RMR: IC (hood) DIT: assumed 10% TEE AEE = TEE-RMR-DIT	1) VM-counts/day 2) Vertical-counts/day	1) § R^2=0.292 2) § R^2=0.27
		CSA (model 7164) uniaxial left hip 7 days, waking hours		1) Counts/day 2) Steps/day	1) § R^2=0.203 2) § R^2=0.176
Bouten *et al.*, 1996 [32]	healthy volunteers n=30 16 male, 14 female 27.1 ±5.0 years 24.1 ±2.3 kg/m²	Tracmor triaxial lower back 7 days, waking hours	TEE: DLW (14 days, first week used) SMR: IC (chamber) AEE = TEE-SMR	Counts/min[g]	§ R^2=0.22 # R^2=0.40

[a] data for study participants extracted from supplement table
[b] analytic sample size was smaller, see crude R² column
[c] extracted data represent VM-acceleration using imputation with wear time at similar time
[d] information obtained from articles' supplement information
[e] no information on Acc type, body position or wear time; assumption in brackets was made based on other references using the same device
[f] SD was calculated from reported standard error (SE) and sample size (SD = SE * √n)
[g] extracted data represent counts/min corrected for transportation vibrations

Acc accelerometer; adj. adjusted; AEE activity related energy expenditure; BMI body mass index; BMR basal metabolic rate; calc calculation; COPD chronic obstructive pulmonary disease; DIT diet-induced thermogenesis; DLW doubly labeled water; FFM fat-free mass; IC indirect calorimetry; n.s. not stated; RMR resting metabolic rate; SD standard deviation; SMR sleeping metabolic rate; TEE total energy expenditure; VM vector magnitude

Table 2 Summary of characteristics derived from the articles in Table 1 concerning study population, accelerometer, and prediction model for all studies and subset of studies with additional prediction model

		All studies (n=28)	Studies with additional prediction model (n=10)[a]
Study population[b]	general population	20	8
	military personnel, soldiers	2	2
	patients	7	0
Sample size	≤ 15	5	1
	16-50	16	6
	51-100	5	1
	> 100	2	2
Sex	male & female	21	6
	female only	6	3
	male only	1	1
Age range	< 40 years	12	6
	> 40 years	10	1
	overall age range (ca. 20-70 years)	6	3
Applied Acc type	uniaxial	12	4
	triaxial	15	6
	biaxial	1[c]	0
Acc recording period	≤ 1 week	12	5
	> 1 week	16	5
Acc body position	trunk (lower back, hip, waist, chest)	20[c]	7
	limbs (wrist, upper arm, ankle, thigh)	8[c]	3
Acc wear time	waking hours	21[c]	6
	24 hours	7	4
Acc output metric	counts/interval	16	6
	steps/interval	6	1
	acceleration	3	2
	other	3	1
Acc output type[d]	uniaxial output	15	6
	triaxial output	13	4
Association/ prediction model	linear	22	9
	non-linear	4	1
	both	2	0

[a] not including Pomeroy *et al.*[17]

[b] for Skipworth *et al.*[28]: allocation to two categories because of various populations (general population and patients)

[c] no information in Adams *et al.*[31] and Colbert *et al.*[16]; made assumptions based on other references using the same device; for Adams: Acc body position=trunk, Acc wear time=waking hours; for Colbert/SenseWear Pro3: Acc type=biaxial, Acc body position=limbs

[d] variable combines information from Acc type and Acc output metric; *steps* output from biaxial or triaxial devices were assigned to “uniaxial output”

Values are number of studies; Acc accelerometer

Table 3 Additional study characteristics and effects of added predictors on total R^2 in subset of studies with prediction model (n=10)

Reference	**Crude-R^2** Acc-output vs. AEE (§) or AEE/kg (#) **Analytic sample size (n)** *(from Table 1)*	**Prediction model** linear (Pearson correlation, regression) nonlinear (Spearman correlation, log linear regression)	**Predictors of final model(s)** (best model in bold)	**Total-R^2** AEE (§) AEE/kg (#) **Partial-R^2** (for Acc-output)	**Other factors** (not included in final model)
Horner *et al.*, 2013 [25]	R^2=0.07 (vs. log AEE) (n=149)	Log linear regression	log counts/day, height, sex (stratified)	§ R^2= 0.41 § Partial-R^2=0.06	weight
Tudor-Locke *et al.*, 2012 [33]	§ R^2=0.0752 (n=54)	Pearson correlation Linear regression	steps/day, peak 30min cadence (steps/min), time in cadence band	§ R^2=0.33	age, height, BMI, weight
Kinnunen *et al.*, 2012 [26]	n.s. (n=22)	Linear regression (stepwise)	normalized hand movement/min, weight, height	# R^2=0.74 **# adj.R^2=0.70** # Partial-R^2=0.41	
van Hees *et al.*, 2011 [19]	§ R^2=0.21 # R^2=0.27[a] (n=48)	Linear regression	**1) VM-acceleration, weight** 2) VM-acceleration, body site, body site* VM-acceleration	**1) § R^2=0.31** 2) § R^2=0.18	height, age, arm length; squared VM-acceleration
	§ R^2=0.09 # R^2=0.05[a] (n=26)		1) VM-acceleration, weight **2) VM-acceleration, body site, body site* VM-acceleration**	1) § R^2=0.05 **2) § R^2=0.19**	
Bonomi *et al.*, 2010 [21]	§ R^2=0.29 # R^2=0.50 (n=30)	Linear regression (stepwise)	1) counts/day, weight **2) counts/day, FFM**	1) § R^2=0.46; Partial-R^2=0.16 **2) § R^2=0.53**; Partial-R^2=0.23	1) height, age, sex 2) FM, age, sex
Bonomi *et al.*, 2009 [22]	n.s. (n=15)	Linear regression (stepwise)	1) counts/day, weight **2) counts/day, FFM**	1) § R^2=0.47; Partial-R^2=0.21 **2) § R^2=0.60**; Partial-R^2=0.38	1) height, age, sex 2) FM, age, sex
Assah *et al.*, 2009 [20]	# R^2=0.14 (n=33)	Linear regression	counts/day + 1) urban/rural 2) age, sex **3) age, sex, urban/rural** 4) age, sex, body fat% 5) age, sex, body fat%, urban/rural	# R^2= 1) 0.21 2) 0.34 **3) 0.40** 4) 0.35 5) 0.40	
Plasqui *et al.*, 2005 [24]	# R^2=0.60 (n=29)	Linear regression	1) VM-counts/day, age, height, weight **2) VM-counts/day, age, FFM, FM** 3) Vertical-counts/day, age, height, weight	1) § R^2=0.81; § Partial-R^2=0.33 **2) § R^2=0.86** 3) § R^2=0.77	
Mâsse *et al.*, 2004 [23]	§ R^2=0.043 (n=136)	Pearson correlation Linear regression	1) counts/day, weight **2) counts/day, FFM**	1) § R^2=0.092; Partial-R^2=0.05 **2) § R^2=0.125**; Partial-R^2=0.04	
Pomeroy *et al.*, 2011 [17,b]	n.s. (n=50)	(adj. Spearman correlation) Linear regression	none (partial for age, sex, height)	§ Partial-R^2=0.22 (Actigraph MTI)	
			none (partial for age, sex, height)	§ Partial-R^2=0.18 (Dynastream AMP-331)	

[a] information obtained from articles' supplement information

[b] this article only reported values for partial R^2 and was therefore not taken into account in analysis and summary tables

Acc accelerometer; adj. adjusted; AEE activity related energy expenditure; BMI body mass index; FFM fat-free mass; FM fat mass; n.s. not stated; VM vector magnitude

Table 4 Summary of characteristics derived from articles in Table 3 concerning additional predictors for studies with additional prediction model

		Studies with additional prediction model (n=10)[a]
Additional predictors[b]	sex	2
	age	2
	height	3
	weight	7
	fat-free mass	4
	body fat	2
	other	4
Maximal number of additional predictors[c]	1	4
	2	4
	3	2
Partial R^2 [c] (for Acc output)	n.s.	4
	0.00-0.20	2
	0.21-0.40	3
	0.41-0.60	1
	0.61-0.80	0
	> 0.80	0

[a] not including Pomeroy *et al.*[17]

[b] allocation to more than one category possible because of various developed models

[c] in studies with various models the best model was selected (defined as having greatest total R^2 along with lowest number of predictors)

Values are number of studies; Acc accelerometer; n.s. not stated

Supplementary Information

Prediction of activity related energy expenditure using accelerometer derived physical activity under free-living conditions – a systematic review

Stephanie Jeran, Astrid Steinbrecher, Tobias Pischon

Corresponding author: Tobias Pischon, Molecular Epidemiology Research Group, Max Delbrück Center for Molecular Medicine (MDC) in the Helmholtz Association, Robert-Rössle-Straße 10, 13125 Berlin, Germany, E-mail: tobias.pischon@mdc-berlin.de, Phone: ++49 30/9406-4563, Fax: ++49 30/9406-4576

Table S1 Characteristics and results of articles on comparison of AEE estimates of accelerometer and DLW derived AEE

Reference	Study participants **Population** **Sample size (n)** **Sex (male/female)** **Age range** (or mean ±SD) **BMI/weight range** (or mean ±SD)	Accelerometer **Device name** **Acc type** **Body position** **Recording period** **Wear time**	Energy expenditure **DLW measurement period** **RMR/BMR/SMR** measurement or calculation (equation) **DIT included** **AEE_{DLW} calculation**	**Estimated Accelerometer Output** AEE_{Acc} calculation Software information	**R^2** AEE_{Acc} vs. AEE_{DLW} **SEE** (derived from Pearson/ Spearman correlation, Linear regression)	**Mean Difference ±SD** AEE_{Acc} - AEE_{DLW} **LOA** (derived from Bland-Altman) **Range of difference** **RMSE**	**Regression of differences coefficient r (p-value)** (derived from Bland-Altman)	**ICC** **CCC** (95%CI)
Silva *et al.*, 2015 [1]	crossover intervention study sample n=17 male 20-38 years 20.2-26.8 kg/m²	Actiheart n.s. (uniaxial)[a] +HR chest 4 days (condition I), 3 days (wash-out), 4 days (condition II), 24 hours	TEE: DLW (2x 4 days, 3 days wash-out) RMR: IC (mask) DIT: 10% TEE AEE = TEE-RMR-DIT	**condition I = Placebo** AEE_{Acc} 1) Acc model 2) Acc+HR-step model 3) Acc+HR-group model Software: 4.0.99	1) R^2=0.17 SEE: 1886 kJ/day 2) R^2=0.66 SEE: 1214 kJ/day 3) R^2=0.50 SEE: 1465 kJ/day	(in kJ/day) 1) Diff: -3315.2 LOA: -6917.6, 287.2 2) Diff: -1146.9 LOA: -3457.3, 1163.5 3) Diff: -572.3 LOA: -4000.3, 2855.7	1) r= -0.61 (p=0.009) 2) r= -0.26 (p=0.314) 3) r=0.27 (p=0.297)	1) CCC=0.10 2) CCC=0.67 3) CCC=0.67
				condition II = Caffeine AEE_{Acc} 1) Acc model 2) Acc+HR-step model 3) Acc+HR-group model Software: 4.0.99	1) R^2=0.39 SEE: 1484 kJ/day 2) R^2=0.64 SEE: 1155 kJ/day 3) R^2=0.51 SEE: 1333 kJ/day	(in kJ/day) 1) Diff: -3557.7 LOA: -6383.5, -731.8 2) Diff: -856.0 LOA: -3487.5, 1775.5 3) Diff: -300.1 LOA: -4038.5, 3438.3	1) r= -0.62 (p=0.008) 2) r=0.26 (p= -0.261) 3) r=0.49 (p= -0.044)	1) CCC=0.13 2) CCC=0.71 3) CCC=0.65
Calabro *et al.*, 2015 [2] [b]	older adults n=29 n.s. 60-78 years n.s.	SenseWear Mini Armband n.s. (triaxial)[a] n.s. (upper arm)[a] 14 days, n.s.	TEE: DLW (n.s.) RMR: measured DIT: 10% TEE AEE = TEE-RMR-DIT	AEE_{Acc}		absolute percent error: 28.4%		

Farooqi *et al.*, 2013 [3]	COPD outpatients n=19 female 59.7-80 years 18.5-30.0 kg/m²	SenseWear Pro2 Armband biaxial right arm triceps 14 days, 24 hours	TEE: DLW (14 days) RMR: IC (hood) AEE = TEE-RMR(IC)	RMR: calc (Harris-Benedict) 1) $AEE_{Acc} = TEE_{Acc}$-RMR(HB) 2) $AEE_{Acc} = TEE_{Acc}$-RMR(IC) Software: 5.1 & 6.1	1) 5.1: R^2=0.35 6.1: R^2=0.42 2) 5.1: R^2=0.50 6.1: R^2=0.56	(in kJ/day) 1) 5.1: Diff: -385 ±686 LOA[c]: -1757, 987 6.1: Diff: -1114 ±634 LOA[c]: -2382, 154 2) 5.1: Diff: 21 ±726 LOA[c]: -1431, 1473 6.1: Diff: -709 ±667 LOA[c]: -2043, 625	1) 5.1: r=0.19 (p=0.44) 6.1: r=0.20 (p=0.41)	1) 5.1: ICC=0.53 (95%: 0.18, 0.79) 6.1: ICC=0.31 (95%: -0.10, 0.69) 2) 5.1: ICC=0.70 (95%: 0.20, 0.89) 6.1: ICC=0.59 (95%: -0.18, 0.85)
		Actiheart uniaxial + HR chest 14 days, 24 hours		1) AEE_{Acc} 2) AEE_{Acc} using RMR(IC)	1) R^2=0.42 2) R^2=0.42	(in kJ/day) 1) Diff: -1128 ±586 LOA[c]: -2300, 44 2) Diff: -709 ±786 LOA[c]: -2281, 863	1) r=0.03 (p=0.91)	1) ICC=0.29 (95%: -0.09, 0.67) 2) ICC=0.55 (95%: -0.18, 0.80)
Slinde *et al.*, 2013 [4]	overweight or obese lactating women n=62 24.6-41.3 years 25.2-37.4 kg/m²	SenseWear Pro2 Armband n.s. (biaxial)[a] right upper arm 8 days, 24 hours	TEE: DLW (14 days) RMR: IC (hood) DIT: 10% TEE AEE = TEE-RMR-DIT	AEE_{Acc} Software: 5.1 & 6.1		(in kcal/day) 5.1: Diff: -11 ±384 range: -950 to 1032 LOA[c]: -779, 757 6.1: Diff: -581 ±266 range: -1348 to -17 LOA[c]: -1113, -49	5.1: r= -0.517 (p<0.001) 6.1: r=0.273 (p=0.032)	
Löf *et al.*, 2013 [5]	healthy volunteers n=20 female 22-45 years 17.7-33.6 kg/m²	Actiheart uniaxial + HR chest 14 days, waking hours	TEE: DLW (14 days) BMR: IC (hood) AEE = TEE-BMR	$AEE_{Acc} = AEE_{Actiheart}$ (Counts+HR individual calibration) + $AEE_{Nonweartime}$ (MET*duration of activties from diary)		(in kJ/day) Diff(AEE): 740 ±1740 LOA[c]: -2200, 3680 (in kJ/kg/day) Diff(AEE/kg): 11.5 ±25.2 LOA[c]: -38.9, 61.9	r=0.42 (p>0.05)	
		RT3 triaxial right hip 14 days, waking hours		$AEE_{Acc} = AEE_{RT3}$ + $AEE_{Nonweartime}$ (MET*duration of activities from diary)		(in kJ/day) Diff(AEE): -2010 ±910 LOA[c]: -3830, -190 (in kJ/kg/day) Diff(AEE/kg): -30.5 ±15.2 LOA[c]: -60.9, -0.1	r=0.11 (p>0.05)	
		IDEEA biaxial thighs, feet, sternum 5 days, waking hours	AEE(5days) = TEE(5days)-BMR	$TEE_{Acc} = TEE_{IDEEA}$ + $EE_{Nonweartime}$ (MET*duration of activties from diary) RMR_{Acc}: EE measured by IDEEA when lying down $AEE_{Acc} = TEE_{Acc} - RMR_{Acc}$		(in kJ/day) Diff(AEE): -1750 ±1325 LOA[c]: -4400, 900 (in kJ/kg/day) Diff(AEE/kg): -26.5 ±23.0 LOA[c]: -72.5, 19.5	r=0.28 (p>0.05)	

Whybrow *et al.*, 2013 [6]	healthy volunteers n=14[d] 7 male, 7 female 20-55 years 25.7 ±5.6 kg/m²	IDEEA n.s. (biaxial)[a] + HR n.s. (thighs, feet, sternum)[a] 14 days, 24 hours	TEE: DLW (12 days free-living +2 days room calorimeter) RMR: IC (hood) DIT: 10% TEE AEE = TEE-RMR-DIT	RMR_{IDEEA}: estimated by IDEEA; RMR_{IC}: IC measurement + sitting/standing routines DIT: 10% of energy intake added to EE_{IDEEA} AEE_{Acc} = EE_{IDEEA}+DIT-RMR 1) using RMR_{IDEEA} 2) using RMR_{IC}	1) R^2=0.414 2) R^2=0.356 (n=12)	(in MJ/day) 1) Diff: 1.7 LOA: -0.5, 3.8 2) Diff: 1.3 LOA: -2.7, 4.5 (n=12)		1) CCC=0.083 2) CCC=0.063 (n=12)
Villars *et al.*, 2012 [7]	healthy volunteers n=35 male 18-55 years 25.2 ±4.0 kg/m²	RT3 triaxial waist 7-10 days, 24 hours	TEE: DLW (10 days) RMR: IC (hood) DIT: 10% TEE AEE/kg = TEE-RMR-DIT**/kg**	AEE_{Acc}**/kg**	R^2=0.22	(in kJ/kg/day) Diff: -19.1 ±21.0 LOA: -61.1; 22.9 RMSE: 28.2	n.s.	ICC=0.40
		Actiheart n.s. (uniaxial + HR)[a] chest 7-10 days, 24 hours		AEE_{Acc}**/kg** 1) from ACC 2) from ACC+HR group calibration model 3) from ACC+HR individual calibration model	1) R^2=0.38 2) R^2=0.38 3) R^2=0.70	(in kJ/kg/day) 1) Diff: -27.3 ±19.0 LOA: -65.2; 10.6 RMSE: 33.1 2) Diff: -7.6 ±20.2 LOA: -47.9; 32.7 RMSE: 21.3 3) Diff: -4.6 ±13.1 LOA: -30.8; 21.6 RMSE: 13.7	1) n.s. 2) r=0.00 3) r=0.41	1) ICC=0.47 2) ICC=0.62 3) ICC=0.81
Kinnunen *et al.*, 2012 [8]	conscripts n=22 male 19-20 years weight 57-111 kg	Polar activity recorder uniaxial non-dominant wrist 7 days, 24 hours	TEE: DLW (7 days) BMR: calc (Schofield) DIT: 10% TEE AEE**/kg** = TEE-BMR-DIT/kg	AEE_{Acc}**/kg** =TEE_{Acc}-BMR-DIT/kg	R^2=0.71 (p<0.001) SEE: 12.5 kJ/kg/day		r= -0.33 (p=0.140)	
Tanhoffer *et al.*, 2012 [9]	people with spinal cord injury, manual wheelchair n=14 13 male, 1 female 18-65 years 19-30 kg/m²	SenseWear Armband biaxial right upper arm 2 days, waking hours	TEE: DLW (14 days) BMR: IC (spirometry metabolic measurement) DIT: 10% TEE AEE = TEE-BMR-DIT	AEE_{Acc} (derived from TEE_{Acc} and METs-system) Software 7.0	R^2=0.16 (p=0.159)	(in kJ/day) Diff: -67 ±2758 LOA: -5472, 5338		

Skipworth *et al.*, 2011 [10]	cancer patients, healthy subjects n=14 12 male, 2 female 25-76 years 20.4-33.5 kg/m²	ActivPAL uniaxial right thigh 14 days, 24 hours	TEE: DLW (14 days) RMR: IC (hood) RMR: calc (Schofield) AEE = TEE-RMR	AEE_{Acc} = total METs/day - nonactivity METs/day	R^2=0.80 (nonlinear)	(in kcal/day) Diff: -18 ±347 LOA: -699, 663		
Löf, 2011 [11]	non-pregnant n=21; women 22-45 years 17.7-33.5 kg/m²	IDEEA biaxial thighs, feet, sternum 5 days, waking hours	TEE: DLW (5 days) BMR: IC (hood) AEE = TEE-BMR	TEE_{Acc}: derived from time spent in 9 IDEEA categories with corresponding MET-value AEE_{Acc}=TEE_{Acc}-BMR	R^2=0.19			
	pregnant n=18; women 24-41 years 18-32 kg/m² (pre-pregnant state)			s. above, MET-values multiplied by 0.88	R^2=0.24			
Mackey *et al.*, 2011 [12]	older volunteers n=19 11 male, 8 female 78-89 years 22.3-34.9 kg/m²	SenseWear Pro Armband biaxial right upper arm 14 days, 24 hours	TEE: DLW (14 days) RMR: IC (hood) DIT: 10% TEE AEE = TEE-RMR-DIT	TEE_{Acc} RMR: calc (Harris-Benedict) AEE_{Acc} =TEE_{Acc}-RMR(HB) Software 5.1 & 6.1	6.1: R^2=0.578 5.1: R^2=0.618	(in kcal/day) 6.1: Diff: -156 ±198 LOA[c]: -552, 240 range: -495 to 403 5.1: Diff: -108 ±185 LOA[c]: -478, 262 range: -478 to 278	6.1: r= -0.054 (p=0.826) 5.1: r= -0.039 (p=0.875)	6.1: ICC=0.645 (95%: 0.146, 0.862) 5.1: ICC=0.720 (95%: 0.356, 0.887)
Colbert *et al.*, 2011 [13]	older adults n=56 12 male, 44 female 74.7 ±6.5 years 25.8 ±4.2 kg/m²	ActiGraph GT1M uniaxial right waist 10 days, waking hours	TEE: DLW (10 days) RMR: IC (hood) DIT: 10% TEE AEE = TEE-RMR-DIT	1) AEE_{Acc} (derived from Freedson equation) 2) AEE_{Acc} (derived from Crouter equation)	1) R^2=0.239 2) R^2=0.360 Spearman correlation	(in kcal/day) 1) Diff: -125 ±209 LOA[c]: -543, 293 2) Diff: 342 ±256 LOA[c]: -170, 854		
		SenseWear Pro3 Armband n.s. (biaxial)[a] n.s. (upper arm)[a] 10 days, waking hours		AEE_{Acc} (derived from TEE_{Acc}) Software 5.12	R^2=0.229 Spearman correlation	(in kcal/day) Diff: -398 ±241 LOA[c]: -880, 84		

Assah *et al.*, 2011 [14]	healthy urban/rural Cameroonians n=33 16 male, 17 female 25-50 years 27.1 ±4.6 kg/m²	Actiheart uniaxial + HR chest 7 days, 24 hours	TEE: DLW (7 days) RMR: IC (handheld) SMR: calc (=0.9*RMR) DIT: 10% TEE AEE = TEE-RMR-SMR-DIT	AEE_{Acc} 1) from Acc 2) from Acc+HR-step individual calibration 3) from Acc+HR-group calibration	1) R^2=0.29 2) R^2=0.16 3) R^2=0.15	(in kJ/kg/day) 1) Diff: -26.6 ±27.0[e] LOA: -79.3, 26.1 RMSE: 37.5 2) Diff: -5.4 ±29.3[e] LOA: -62.7, 51.9 RMSE: 29.3 3) Diff: -9.1 ±28.7[e] LOA: -65.8, 47.6 RMSE: 29.9	
Johannsen *et al.*, 2010 [15]	volunteers n=30 15 male, 15 female 24-60 years 17.8-31.5 kg/m²	SenseWear Pro3 Armband biaxial right upper arm 14 days, 24 hours	TEE: DLW (14 days) RMR: calc (WHO) DIT: 10% TEE AEE = TEE-RMR-DIT	AEE_{Acc} =TEE_{Acc}-RMR-DIT Software 6.1	R^2=0.51 (p<0.001)	(in kcal/day) Diff: -123 ±278 LOA[c]: -679, 433	ICC=0.63 (95% 0.47, 0.77)
		SenseWear Mini Armband triaxial left upper arm 14 days, 24 hours		AEE_{Acc} =TEE_{Acc}-RMR-DIT Software 7.0	R^2=0.48 (p<0.001)	(in kcal/day) Diff: -119 ±286 LOA[c]: -691, 453	ICC=0.63 (95% 0.47, 0.77)
Assah *et al.*, 2009 [16]	healthy urban/rural Cameroonians n=33 16 male, 17 female 25-50 years 27.1 ±4.6 kg/m²	ActiGraph GT1M uniaxial waist 7 days, 24 hours	TEE: DLW (7 days) RMR: IC (handheld calorimeter) SMR: calc (=0.9*RMR) (assuming 8h sleep) DIT: 10% TEE AEE=TEE-RMR-SMR-DIT	AEE_{Acc} 1) from Freedson 2) from Hendelman 3) from Swartz each with **a) RMR: calc Schofield** **b) RMR: 1 METs**		(in kJ/kg/day) 1)a) Diff: -6.29 ±31.94[e] LOA: -68.85, 56.27 RMSE: 32.06 b) Diff: -20.66 ±28.61[e] LOA: -76.70, 35.38 RMSE: 34.92 2)a) Diff: 23.45 ±36.42[e] LOA: -47.95, 94.86 RMSE: 42.86 b) Diff: 9.09 ±31.71[e] LOA: -53.11, 71.28 RMSE: 32.54 3)a) Diff: 23.28 ±34.93[e] LOA: -45.15, 91.71 RMSE: 41.52 b) Diff: 8.91 ±30.73[e] LOA: -51.33, 69.15 RMSE: 31.55	

Maddison *et al.*, 2009 [17] [b]	n.s., n=36 16 male, 20 female 18-56 years weight 75.9 ±14.8 kg	RT3 triaxial n.s. 14 days, n.s.	TEE: DLW (14 days) RMR: n.s. DIT: n.s. (10% TEE) AEE = TEE-RMR-DIT	AEE_{Acc}		Diff: -485 kJ (15% underestimation)	
Yamada *et al.*, 2009 [18]	healthy elderly n=32 14 male, 18 female 64-87 years 17.0-27.5 kg/m²	Lifecorder uniaxial back of waist 14 days, waking hours	TEE: DLW (14 days) RMR: IC SMR: calc (=0.95*RMR) DIT: 10% TEE AEE=TEE-DIT-RMR-SMR	AEE_{Acc}		(in MJ/day) Diff: -0.91 range of LOA: 3.84 LOA[c]: -2.83, 1.01	
		n.s. (Actimarker)[a] triaxial back of waist 14 days, waking hours		AEE_{Acc}		(in MJ/day) Diff: 0.03 range of LOA: 3.03 LOA[c]: -1.49, 1.54	
Jacobi *et al.*, 2007 [19]	healthy overweight/obese n=13[d] female 38.3 ±10.5 years 34.2 ±6.4 kg/m²	RT3 triaxial waistline above hip 14 days, waking hours	TEE: DLW (14 days) RMR: IC (hood) DIT: 10% TEE AEE = TEE-RMR-DIT	RMR: calc (Harris Benedict) 1) AEE_{Acc} 2) AEE_{Acc}-corrected= AEE_{Acc}/RMR(HB)*RMR (IC)	1) R^2=0.30 2) R^2=0.45	(in kcal/day) Diff[c]: -120 ±265 LOA: -385, 145 rel. Diff: -17.1 ±16.7 %	
		TriTrac-R3D triaxial waistline above hip 14 days, waking hours		RMR: calc (Harris Benedict) AEE_{Acc}-corrected= AEE_{Acc}/RMR(HB)*RMR (IC)	R^2=0.13 (n=12)	(in kcal/day) Diff[c]: -148.5 ±738.5 LOA: -887, 590 rel. Diff: -20.0 ±44.6 % (n=12)	
St-Onge *et al.*, 2007 [20]	healthy & diabetics n=45[d] 13 male, 32 female 20.1-78.2 years 17.9-34.3 kg/m²	SenseWear Armband biaxial right triceps 10 days, 24 hours	TEE: DLW (10 days) RMR: IC (hood) DIT: 10% TEE AEE =TEE-RMR-DIT	TEE_{Acc}, RMR_{Acc} DIT_{Acc}: 10% of TEE_{Acc} AEE_{Acc}=TEE_{Acc}-DIT_{Acc}-RMR_{Acc} Software 4.02	R^2=0.49 SEE: ±179 kcal/day (n=41)	Diff: -225 kcal/day (n=41)	ICC=0.46 (95%: 0.19, 0.67; p<0.01) (n=41)
Johansson *et al.*, 2006 [21]	subsample validation study n=8 6 male, 2 female 28-63 years 21.2-36.0 kg/m²	ActiGraph MTI (model 7164) uniaxial lower back 14 days, waking hours	TEE: DLW (14 days) RMR: IC (Douglas bags, Spirometer) DIT: 10% TEE AEE**/kg** = TEE-RMR-DIT/kg	AEE_{Acc}**/kg**		(in kJ/kg/day) Diff: 0.0 95%CI: -30.5; 30.4 LOA: -71.5; 71.4 RMSE: 34.1	

Leenders *et al.*, 2001 [22]	healthy volunteers n=13 female 21-37 years 19.9-27.7 kg/m²	Tritrac-R3D triaxial right hip 7 days, waking hours	TEE: DLW (7 days) RMR: IC (hood) DIT: 10% TEE AEE = TEE-RMR-DIT	AEE_{Acc}		(in kcal/day) Diff: -320 range: -780 to 89 (35% underestimation of DLW)
		CSA (model 7164) uniaxial left hip 7 days, waking hours		AEE_{Acc} [f]		(in kcal/day) Diff: -495 range: -824 to 5 (59% underestimation of DLW)
Starling *et al.*, 1999 [23]	healthy elderly n=67 32 male, 35 female 45-84 years male: 25.7 ±4.5 kg/m²; female: 24.8 ±3.9 kg/m²	Caltrac uniaxial hip 9 days, waking hours	TEE: DLW (10 days) RMR: IC (hood) DIT: 10% TEE AEE = TEE-RMR-DIT	AEE_{Acc}		(in kcal/day) <u>women:</u> Diff[c]: -494 LOA: -911, -75 <u>men:</u> Diff[c]: -657 LOA: -1408, 96
Gardner *and* Poehlman, 1998 [24]	PAOD patients n=22 20 male, 2 female 53-88 years 20.4-41.2 kg/m²	Caltrac uniaxial hip 2 weekdays, waking hours	TEE: DLW (10 days) RMR: IC (hood) DIT: 10% TEE AEE = TEE-RMR-DIT	AEE_{Acc}	R^2=0.696 SEE: 77 kcal/day	Diff: 117 kcal/day

[a] no information on device name, Acc type, body position or wear time; assumption in brackets based on other references using the same accelerometer
[b] article was only available as abstract
[c] converted from reported results: LOA = mean difference ± 2*SD; Diff = upper LOA limit - (LOA range/2); SD = LOA range/2
[d] analytic sample size was smaller, see results columns
[e] SD was calculated from reported standard error (SE) and sample size (SD = SE * √n)
[f] AEE calculation equation stated in article

Acc accelerometer; AEE activity related energy expenditure; BMI body mass index; BMR basal metabolic rate; calc calculated/calculation; CCC concordance correlation coefficient; COPD chronic obstructive pulmonary disease; Diff difference; DIT diet induced thermogenesis; DLW doubly labeled water; EE energy expenditure; HB Harris-Benedict-equation; HR heart rate; IC indirect calorimetry; ICC intraclass correlation coefficient; LOA limits of agreement; MET Metabolic equivalent of task; n.s. not stated; PAOD peripheral arterial occlusive disease; R² coefficient of determination, squared correlation coefficient; RMR resting metabolic rate; RMSE root mean squared error; SD standard deviation; SEE/SE standard (estimated) error; SMR sleeping metabolic rate; 95%CI 95% confidence interval

References

1. Silva AM, Santos DA, Matias CN, Judice PB, Magalhaes JP, Ekelund U *et al.* Accuracy of a combined heart rate and motion sensor for assessing energy expenditure in free-living adults during a double-blind crossover caffeine trial using doubly labeled water as the reference method. *Eur J Clin Nutr* 2015; **69**: 20-27.
2. Calabro MA, Kim Y, Franke WD, Stewart JM, Welk GJ. Objective and subjective measurement of energy expenditure in older adults: a doubly labeled water study. *Eur J Clin Nutr* 2015; **69**: 850-855.
3. Farooqi N, Slinde F, Haglin L, Sandstrom T. Validation of SenseWear Armband and ActiHeart monitors for assessments of daily energy expenditure in free-living women with chronic obstructive pulmonary disease. *Physiol Rep* 2013; **1**: e00150.
4. Slinde F, Bertz F, Winkvist A, Ellegard L, Olausson H, Brekke HK. Energy expenditure by multisensor armband in overweight and obese lactating women validated by doubly labeled water. *Obesity (Silver Spring)* 2013; **21**: 2231-2235.
5. Lof M, Henriksson H, Forsum E. Evaluations of Actiheart, IDEEA(R) and RT3 monitors for estimating activity energy expenditure in free-living women. *J Nutr Sci* 2013; **2**: e31.
6. Whybrow S, Ritz P, Horgan GW, Stubbs RJ. An evaluation of the IDEEA activity monitor for estimating energy expenditure. *Br J Nutr* 2013; **109**: 173-183.
7. Villars C, Bergouignan A, Dugas J, Antoun E, Schoeller DA, Roth H *et al.* Validity of combining heart rate and uniaxial acceleration to measure free-living physical activity energy expenditure in young men. *J Appl Physiol (1985)* 2012; **113**: 1763-1771.
8. Kinnunen H, Tanskanen M, Kyrolainen H, Westerterp KR. Wrist-worn accelerometers in assessment of energy expenditure during intensive training. *Physiol Meas* 2012; **33**: 1841-1854.
9. Tanhoffer RA, Tanhoffer AI, Raymond J, Hills AP, Davis GM. Comparison of methods to assess energy expenditure and physical activity in people with spinal cord injury. *J Spinal Cord Med* 2012; **35**: 35-45.
10. Skipworth RJ, Stene GB, Dahele M, Hendry PO, Small AC, Blum D *et al.* Patient-focused endpoints in advanced cancer: criterion-based validation of accelerometer-based activity monitoring. *Clin Nutr* 2011; **30**: 812-821.
11. Lof M. Physical activity pattern and activity energy expenditure in healthy pregnant and non-pregnant Swedish women. *Eur J Clin Nutr* 2011; **65**: 1295-1301.
12. Mackey DC, Manini TM, Schoeller DA, Koster A, Glynn NW, Goodpaster BH *et al.* Validation of an armband to measure daily energy expenditure in older adults. *J Gerontol A Biol Sci Med Sci* 2011; **66**: 1108-1113.
13. Colbert LH, Matthews CE, Havighurst TC, Kim K, Schoeller DA. Comparative validity of physical activity measures in older adults. *Med Sci Sports Exerc* 2011; **43**: 867-876.
14. Assah FK, Ekelund U, Brage S, Wright A, Mbanya JC, Wareham NJ. Accuracy and validity of a combined heart rate and motion sensor for the measurement of free-living physical activity energy expenditure in adults in Cameroon. *Int J Epidemiol* 2011; **40**: 112-120.
15. Johannsen DL, Calabro MA, Stewart J, Franke W, Rood JC, Welk GJ. Accuracy of armband monitors for measuring daily energy expenditure in healthy adults. *Med Sci Sports Exerc* 2010; **42**: 2134-2140.
16. Assah FK, Ekelund U, Brage S, Corder K, Wright A, Mbanya JC *et al.* Predicting physical activity energy expenditure using accelerometry in adults from sub-Sahara Africa. *Obesity (Silver Spring)* 2009; **17**: 1588-1595.
17. Maddison R, Jiang Y, Hoorn SV, Mhurchu CN, Lawes CM, Rodgers A *et al.* Estimating energy expenditure with the RT3 triaxial accelerometer. *Res Q Exerc Sport* 2009; **80**: 249-256.

18. Yamada Y, Yokoyama K, Noriyasu R, Osaki T, Adachi T, Itoi A *et al.* Light-intensity activities are important for estimating physical activity energy expenditure using uniaxial and triaxial accelerometers. *Eur J Appl Physiol* 2009; **105**: 141-152.
19. Jacobi D, Perrin AE, Grosman N, Dore MF, Normand S, Oppert JM *et al.* Physical activity-related energy expenditure with the RT3 and TriTrac accelerometers in overweight adults. *Obesity (Silver Spring)* 2007; **15**: 950-956.
20. St-Onge M, Mignault D, Allison DB, Rabasa-Lhoret R. Evaluation of a portable device to measure daily energy expenditure in free-living adults. *Am J Clin Nutr* 2007; **85**: 742-749.
21. Patrik Johansson H, Rossander-Hulthen L, Slinde F, Ekblom B. Accelerometry combined with heart rate telemetry in the assessment of total energy expenditure. *Br J Nutr* 2006; **95**: 631-639.
22. Leenders NY, Sherman WM, Nagaraja HN, Kien CL. Evaluation of methods to assess physical activity in free-living conditions. *Med Sci Sports Exerc* 2001; **33**: 1233-1240.
23. Starling RD, Matthews DE, Ades PA, Poehlman ET. Assessment of physical activity in older individuals: a doubly labeled water study. *J Appl Physiol (1985)* 1999; **86**: 2090-2096.
24. Gardner AW, Poehlman ET. Assessment of free-living daily physical activity in older claudicants: validation against the doubly labeled water technique. *J Gerontol A Biol Sci Med Sci* 1998; **53**: M275-280.