# Accepted Manuscript


Title: Validation of the SenseWear Mini activity monitor in 5-12 year-old children

Author: Christiana M.T. van Loo Anthony D. Okely Marijka J. Batterham Trina Hinkley Ulf Ekelund Søren Brage John J. Reilly Gregory E. Peoples Rachel Jones Xanne Janssen Dylan P. Cliff

PII: S1440-2440(16)30057-3
DOI: http://dx.doi.org/doi:10.1016/j.jsams.2016.04.010
Reference: JSAMS 1322

To appear in: *Journal of Science and Medicine in Sport*

Received date: 21-9-2015
Revised date: 8-3-2016
Accepted date: 30-4-2016

Please cite this article as: van Loo CMT, Okely AD, Batterham MJ, Hinkley T, Ekelund U, Brage S, Reilly JJ, Peoples GE, Jones R, Janssen X, Cliff DP, Validation of the SenseWear Mini activity monitor in 5-12 year-old children, *Journal of Science and Medicine in Sport* (2016), http://dx.doi.org/10.1016/j.jsams.2016.04.010

# Validation of the SenseWear Mini activity monitor in 5-12 year-old children

Christiana M.T. van Loo[a], Anthony D. Okely[a], Marijka J. Batterham[b], Trina Hinkley[c], Ulf Ekelund[d,e], Søren Brage[e], John J. Reilly[f], Gregory E. Peoples[g], Rachel Jones[a], Xanne Janssen[f], Dylan P. Cliff[a]

[a]Early Start Research Institute, Faculty of Social Sciences, University of Wollongong, Australia

[b]School of Mathematics and Applied Statistics, University of Wollongong, Australia

[c]Deakin University, Centre for Physical Activity and Nutrition Research, Australia

[d]Norwegian School of Sports Sciences, Norway

[e]MRC Epidemiology Unit, University of Cambridge, United Kingdom

[f]School of Psychological Sciences and Health, University of Strathclyde, Scotland

[g]School of Medicine, University of Wollongong, Australia

Corresponding author: C.M.T. van Loo

E-mail address: cmtvl646@uowmail.edu.au

Word count: 2987

Abstract word count: 204

Number of tables: 1

Number of figures: 2

# Validation of the SenseWear Mini activity monitor in 5-12 year-old children


## Abstract

Objectives: This study aimed to validate SenseWear Mini software algorithm versions 2.2 (SW2.2) and 5.2 (SW5.2) for estimating energy expenditure (EE) in children.

Design: Laboratory-based validation study.

Methods: 57 children aged 5-12 y completed a protocol involving 15 semi-structured sedentary (SED), light-intensity (LPA), and moderate- to vigorous-intensity (MVPA) physical activities. EE was estimated using portable indirect calorimetry (IC). The accuracy of EE estimates (kcal·min$^{-1}$) from SW2.2 and SW5.2 were examined at the group level and individual level using the mean absolute percentage error (MAPE), Bland-Altman plots and equivalence testing.

Results: MAPE values were lower for SW5.2 (30.1% ± 10.7%) than for SW2.2 (44.0% ± 6.2%). Although mean differences for SW5.2 were smaller than for SW2.2 during SED (-0.23 ± 0.22 vs. -0.61 ± 0.20 kcal·min$^{-1}$), LPA (-0.69 ± 0.76 vs. -1.07 ± 0.46 kcal·min$^{-1}$) and MVPA (-2.22 ± 1.15 vs. -2.57 ± 1.15 kcal·min$^{-1}$), limits of agreement did not decrease for the updated algorithms. For all activities, SW2.2 and SW5.2 were not equivalent to IC ($p>0.05$). Errors increased with increasing intensity.

Conclusion: The current SenseWear Mini algorithms SW5.2 underestimated EE. The overall improved accuracy for SW5.2 was not accompanied with improved accuracy at the individual level and EE estimates were not equivalent to IC.



## Keywords

Energy expenditure, physical activity, accelerometry, calorimetry, validation study

## Introduction

Physical activity (PA) is an established determinant of children's health[1] and the energy expenditure (EE) from PA might be particularly important for obesity and chronic disease prevention.[2] Prevalence data show low levels of PA among school-aged children and adolescents,[3-5] making it essential to further understand and promote PA among these age groups. Accurate measures are of critical importance to identify the prevalence of participation in PA, to establish associations with health outcomes, identify correlates of PA, and to evaluate the effectiveness of interventions to promote PA and increase EE.[6] Accelerometer has become the method of choice for objectively measuring habitual PA in children.[7,8] Traditional accelerometers and single-regression equation data reduction approaches typically provide accurate assessments of EE for a limited number of activities. However, the assessment of EE is not accurate over the wide range of lifestyle activities in which children typically participate.[9,10] This is partly due to the biomechanical variation of different activity types and the variability in activity energy costs due to growth and maturation.[11]

Multi-sensor activity monitors could possibly overcome these limitations, and have the potential to make substantial improvements in the measurement of PA and EE during free-living lifestyle behaviours among children. The SenseWear Mini (BodyMedia Inc., Pittsburgh, PA, USA) is a device that combines accelerometry data and multiple physiological signals i.e. heat flux, skin temperature, near-body ambient temperature and galvanic skin response (GSR), using a pattern-recognition-based analysis approach.[12] The arm-mounted SenseWear Mini with integrated physiological sensors has the potential to assess EE of non-ambulatory activities more accurately than traditional accelerometers, especially those worn on the hip. A unique characteristic of the SenseWear activity monitor is that the company continually updates the algorithms as new data become available and are integrated into its pattern recognition system.

Consistent improvements in the estimation of EE using updated data processing algorithms (v.2.0, 2.2 and 5.0) have been found in laboratory and free-living studies in children.[12-14] A recent study by Lee et al.[15] confirmed an improved activity specific accuracy of SenseWear Mini's updated child algorithms (v.5.2; hereafter SW5.2), compared to the previous version (v.2.2; hereafter SW2.2). An ecological design was used to simulate real-world conditions by selecting 12 activities from a larger pool of 24, which were completed in a random order. Although this approach was a strength of the study, it resulted in a small sample size (n<20) for 9 activities, and girls were under-represented (24.4% of the sample). No studies have validated the new algorithms in children <7 y. To date, validation studies have used dependent sample tests to examine differences between previous and updated software versions. However, no studies have investigated whether the EE estimates lie within an acceptable range from the criterion measure. Traditional analyses that fail to reject the null hypothesis of similarity do not necessarily demonstrate that the software algorithms meet an acceptable level of accuracy. Therefore, equivalence testing, where the null hypothesis is reversed to examine the equivalence of two methods, is recommended for validation studies as an alternative approach.[16,17] This study aimed to compare the accuracy of SW2.2 and SW5.2 in school-aged children, during a range of ambulatory and lifestyle activities, by combining standard analyses of measurement agreement with formal testing of equivalence.

**Methods**

Children aged 5-12 y who were without physical or health conditions that would affect their EE or participation in PA were recruited as part of an activity monitor validation study. Participants were required to visit the laboratory twice within a 2- to 4-wk period. The study was approved by the University of Wollongong Health and Medical Human Research Ethics Committee. Parental consent and participant assent were obtained prior to participation.

Participants fasted for 2 hr prior to each laboratory visit. Anthropometric measures were completed using standardised procedures during the first visit while children were wearing light clothing and with shoes removed. BMI (kg/m$^2$) and weight status were calculated.[18] At each visit children were fitted with a SenseWear Mini and a portable respiratory gas analysis system (MetaMax® 3B, Cortex, Biophysics, Leipzig, Germany). Children completed a protocol of 15 semi-structured activities

(Supplementary Table 1), ranging in intensity from sedentary to vigorous. Activities were equally divided over 2 visits and completed in a structured order of increasing intensity for 5 min, except for lying down (10 min). The activity protocol was developed to align with best practice recommendations[19] and included several activities that have been used in previous validation and calibration studies.[9,15] For descriptive purposes, the activities were categorised as sedentary (SED: $<1.5$ METs), light- (LPA: $\geq 1.5$ to $<3$ METs), moderate- (MPA: $\geq 3$ to $<6$ METs) or vigorous-intensity (VPA: $\geq 6$ METs) physical activities based on the Compendium of Energy Expenditure for Youth.[20] Measured and estimated EE values are presented in Supplementary Table 2.

The SenseWear Mini was placed over the triceps muscle of the left arm, according to the company's guidelines. SenseWear Professional Software v.7.0 (SW2.2) and v.8.0 (SW5.2) were used to reduce the data. Accelerometry and additional physiological data combined with personal characteristics such as weight, height, age and sex are integrated in a proprietary algorithm to estimate EE. The analysis of the pattern of signals from the sensors is automatically performed by the movement-specific algorithms and outcomes of EE are exported at 1 min intervals.

Oxygen consumption ($O_2$) and carbon dioxide production ($CO_2$) were assessed using the MetaMax® 3B portable breath-by-breath respiratory gas analysis system to provide the criterion assessment of EE. The participants wore a facemask (Hans Rudolph, Kansas City, MO) covering their nose and mouth, which was held in place by a head harness. Prior to every measurement, the analyser was calibrated according to the manufacturer's guidelines. Breath-by-breath data from IC were downloaded and exported using MetaSoft (version 4.3.2). Mean volume of $O_2$ uptake and $CO_2$ production were converted into units of EE ($kcal \cdot min^{-1}$) using the Weir equation.[21]

The SenseWear Mini and IC were synchronised with an internal computer clock. Data from both SW2.2 and SW5.2 algorithms were compared with indirect calorimetry (IC) to examine whether the new child prediction equation was more accurate for assessing EE. Customised software was used to calculate minute-by-minute EE values and align the outcomes with the Sensewear Mini data.

Mean absolute percentage error (MAPE) and Bland-Altman plots[22] were used to evaluate measurement agreement, individual variability, and systematic bias across the range of activities. MAPE values were calculated as the average of the absolute difference between the software algorithm and IC divided by IC, multiplied by 100%. Pearson correlations were used to evaluate the influence of age and BMI percentile on the performance of SW2.2 and SW5.2. Overall agreement of SenseWear Mini algorithms and IC was determined using the 95% equivalence test. In order to reject the null hypothesis, the 90% confidence intervals (CI; 100%-2α) of SW2.2 or SW5.2 should lie entirely within the predefined equivalence region of ± 10% of the mean for IC. A mixed model ANOVA was used to compute 90% CIs including participants as a random effect to account for repeated measures. Normality tests showed that EE values were skewed. Log transformation was used as Ln(x+1) to meet the assumptions of normal distribution for performing equivalence testing.

**Results**

Descriptive characteristics of the 57 participating children are presented in Table 1. All participants completed the protocol. Data from one child were entirely excluded from the analyses and data from 3 participants for a total of 8 activities were excluded because of IC failure. Minute-by-minute data were partly excluded when aligning IC with SenseWear Mini data, due to activities that were not completed parallel to the 1 min samples of the SenseWear Mini. A total of 4440 minutes were included for analysis, accounting for 98.8% of the total data. All individual activities yielded smaller MAPE values (Figure 1) for SW5.2 (30.1% ± 10.7%) than for SW2.2 (44.0% ± 6.2%). Smallest MAPE values were found in ambulatory activities (slow walk: 32.5%; brisk walk: 34.8% and running: 35.6%) for SW2.2 and in sedentary activities (TV: 13.8%; lying down: 14.7%; computer game: 17.3%; and writing/colouring: 23.9%) for SW5.2. MAPE values for SW2.2 were greater during SED (47.9% ± 2.2%) than during LPA (40.2% ± 6.9%) and MVPA (43.4% ± 7.0%). MAPE values for SW5.2 yielded 19.0% ± 5.2%, 32.6% ± 10.2% and 37.6% ± 6.3% for SED, LPA and MVPA, respectively. Largest relative percentage improvement was found for SED (60.4%). Reasonable improvement was found for LPA (19.0%) and MVPA (13.2%), particularly for slow walk (24.6%), dancing (33.2%) and brisk walk (21.5%). Although clear improvement was shown for all activities, MAPE values for SW5.2 increased with increasing intensity of activity. Furthermore, MAPE values seemed negatively related to age (SW2.2: r= -0.76, $p<0.01$; SW5.2: r = -0.53, $p<0.01$) and BMI percentile (SW2.2: r= -0.37, $p<0.01$; SW5.2: r = -0.32, $p<0.05$).

Bland-Altman plots (Supplementary Figure 1) showed consistent underestimation of EE for both algorithms, although mean differences between the criterion measure and the algorithms for SW5.2 were smaller compared to SW2.2 during SED (-0.23 kcal·min$^{-1}$ vs. -0.61 kcal·min$^{-1}$, respectively), LPA (-0.69 kcal·min$^{-1}$ vs. -1.07 kcal·min$^{-1}$, respectively) and MVPA (-2.22 kcal·min$^{-1}$ vs. -2.57 kcal·min$^{-1}$, respectively). No improvements were detected in 95% limits of agreement (LoA). Random error, defined as the SD of the residuals, was larger for SW5.2 compared to SW2.2 in SED (0.22 kcal·min$^{-1}$ vs. 0.20 kcal·min$^{-1}$, respectively) and LPA (0.76 kcal·min$^{-1}$ vs. 0.46 kcal·min$^{-1}$, respectively), whereas random error for MVPA remained equal (1.15 kcal·min$^{-1}$). Slopes of the regression model were significantly different from zero ($p<0.01$) in all cases. As the difference between algorithms and IC were dependent on average EE estimates, systematic bias was present. Neither SW2.2 nor SW5.2 was equivalent to IC for all activities ($p>0.05$) as none of the 90% CIs were entirely included in the equivalence region (Figure 2). 90% CIs for SW5.2 lay closer to the equivalence zone than for SW2.2, especially for all sedentary activities, slow walk and brisk walk. Means and/or 90% CIs partly overlapped with the equivalence region for lying down, TV, computer game and dancing. The plot shows greater error with increasing intensity for SW5.2.

## Discussion

This study examined the validity of the most recently released SenseWear Mini algorithms for estimating EE in children. The updated algorithms SW5.2 underestimated EE, although overall improved agreement was found at the group level compared to SW2.2, particularly for sedentary activities and some light activities. However, large random error was present at the individual level and none of the estimates were found to be equivalent to the criterion measure for all activities.

The results are broadly in agreement with other SenseWear validation studies, showing a consistent improvement when directly comparing previous with updated algorithms. Improved accuracy for the updated set of child algorithms (v.5.0) was found in a study using doubly labelled water (DLW) as the criterion measure among free-living 10-16 year-olds.[14] Large random error indicated the need for further evaluation at the individual level, and it was unclear if this error differed by the intensity of the activity. Lee et al.[15] included 45 children aged 7-13 y, who wore a portable IC system and a SenseWear Mini while completing 12 randomly selected activities. MAPE values of 17.1% and 4.6% showed overall improvement for SW5.2 during sedentary and light activities, respectively. Although MAPE values for SW5.2 during sedentary activities (19.0%) in our study were similar to those reported by Lee et al.[15] the mean error for light activities (32.6%) was considerably higher. These authors found that SW5.2 was accurate for estimating EE during overground walking-based activities (MAPE for brisk walking: 0.51%; walking at casual pace: 1.91%; slow walking 4.23%). However, ambulatory activities in our protocol revealed larger MAPE values (slow walk: 24.5%; brisk walk: 27.4%). Activities requiring vigorous arm-movements were discussed by Lee et al.[15] because lower MAPE values were detected for SW2.2 compared to SW5.2, indicating that the new algorithm might negatively affect estimates of EE when more upper body movement is involved. All activities in the present study showed smaller MAPE values for SW5.2 compared to SW2.2. In addition, activities with the least upper-body movement yielded low relative percentage improvements (standing class activity: 4.2%; soccer: 6.4%; running: 7.2%) for the new algorithms, whereas activities with more upper body movement yielded higher improvement (basketball, 10.7%; getting ready for school, 19.1%; tidy up, 19.8%; dancing, 33.2%). Based on these findings, it can be suggested that the estimates of EE might be affected during lifestyle activities involving a range of complex activity patterns, rather than the requirement of vigorous arm movements alone. It should be noted that MAPE values were negatively correlated with age and BMI percentile, although the associations were weaker with SW5.2. Thus the algorithms might be less accurate in younger children and those with a lower BMI for their age and sex. This should be considered when applying the assessments in children. The characteristics of the algorithm development samples are unknown, but if the algorithms were developed in older and heavier children, this may have contributed to these findings.

Overall errors were smaller for SW5.2 compared to SW2.2, although LoAs did not decrease. Lee et al.[15] also reported better overall agreement for the new algorithms, however their narrower LoAs were in contrast with our findings. Even though errors increased with increasing intensity in both studies, no systematic bias was reported by Lee et al.[15] Differences in findings could be explained by the different activities included in the protocols or the inclusion of a slightly younger age group and equal numbers of boys and girls in the current study. Furthermore, Lee et al.'s[15] ecological design resulted in a small sample size for some activities. Although all participants completed all activities in our study, fewer overweight and no obese children were included. While a clear reason for the different findings might be hard to establish, it should be noted that conclusions about the accuracy of the updated SW5.2 algorithms should be considered with caution.

Our findings from Bland-Altman plots were similar to those of Calabro et al.,[14] indicating that improved accuracy at the group level with the updated algorithms was not accompanied with improvements at the individual level. LoAs in our plots became notably wider for LPA. This is likely explained by a group of extreme errors for the activities of getting ready for school and dancing. Most of these errors originated from data in overweight children and suggested large overestimation in these particular cases. A study by Bäcklund et al.[23] showed that a previous set of algorithms (v2.0) was more accurate for estimates of EE than the updated SW2.2 in overweight and obese free-living children. A significant underestimation of 18% was detected when the update was applied. The difference between algorithms was particularly high during LPA when directly compared with each other. A correction for overweight and obese children was the company's key focus when updating to algorithms version 5,[14] which might have a negative effect at the individual level for this category and a shift toward overestimation of energy levels might occur.

Despite the improvements for the new algorithms in both previous studies and the current study, overall MAPE values for SW5.2 remain large and non-equivalence between SW5.2 and the criterion measure IC was demonstrated by this study. 90% CIs for sedentary and overground walking (slow walk and brisk walk) lay very close to the equivalence range, indicating that estimates were reasonably accurate for these activities. However, as demonstrated by Bland-Altman plots in Lee et al.'s[15] study and the current study, the equivalence plot confirms that errors increased with increasing intensity for SW5.2. An underestimation (MAPE) of 37.6% for MVPA means that if a 10 year-old boy used 225 kcal during 30min of soccer, SW5.2 would underestimate his EE by 84.6 kcal, which is two times his resting EE (measured EE while lying down) over the same amount of time.

A strength of this study is the large sample size including a broad age range and an equal distribution of age and sex across the sample. Furthermore, the protocol involved a wide range of semi-structured lifestyle activities to assist with generalising the findings to free-living conditions. By evaluating the activity-specific accuracy of the SW2.2 and SW5.2 algorithms at the individual level, we were able to provide insight into measurement errors identified in the previous free-living study.[14] A unique strength of this study was the analysis of equivalence that provides new information to the findings from previous studies showing significantly lower errors for the updated algorithms. By using the equivalence test as an alternative method, we were able to examine whether the reduced measurement errors lay within a conventional range of ±10% of the criterion. It is recommended for future validation studies to use similar methods of analysis, in an effort to directly compare findings. As a potential limitation of this type of testing, it should be noted that although the ±10% is conventional, it is unclear if it represents a clinically meaningful range. Another limitation of this study is that we did not include cycling, an activity that is proven to be difficult to assess with traditional accelerometry-based activity monitors. Furthermore, because the company does not provide detailed information about the proprietary algorithms, it is impossible to independently evaluate how the algorithms might affect the outcomes. Future validation research should also focus on the accuracy of new algorithms in obese children.

**Conclusion**

The SW5.2 algorithms demonstrated improved accuracy at the group level, particularly for sedentary and ambulatory activities, however measurement errors remain large and estimates of EE were not found to be equivalent to IC. At the individual level, systematic bias was found for both algorithms and errors increased with increasing intensity for SW5.2.

## Practical Implications

- Updated SenseWear Mini software algorithms should be used for improved assessment of EE in children.
- Outcomes from the software algorithms should be interpreted with caution, particularly for individual values rather than for groups of children, and for high intensity activities.
- Equivalence testing combined with other tests of agreement should be used in future validation studies to directly compare findings and provide insight into the clinical acceptance of measurement errors.

## Acknowledgments

We would like to thank all children and their parents for their participation. We also thank Melinda Smith for her assistance with recruitment and data collection. This study was funded by the National Heart Foundation of Australia (G11S5975).

**Figure 1.** Mean absolute percentage error of algorithms version 2.2 (SW2.2) and 5.2 (SW5.2) relative to the criterion measure portable indirect calorimetry across all the activities.

**Figure 2.** 95% equivalence test for logarithmically transformed energy expenditure data across sedentary (SED), light- (LPA) and moderate- to vigorous-intensity (MVPA) physical activities. Methods are equivalent if 90% confidence intervals lie entirely within the equivalence region of IC. *IC, indirect calorimetry; SW2.2, algorithms version 2.2; SW5.2, algorithms version 5.2.

## References

1. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. *Int J Behav Nutr Phys Act* 2010; 7:40-56.

2. Wareham NJ, van Sluijs EMF, Ekelund U. Physical activity and obesity prevention: a review of the current evidence. *Proc Nutr Soc* 2005; 64(2):229-247.

3. Australian Bureau of Statistics. Australian Bureau of Statistics, Australian Health Survey: Physical Activity. Canberra 2011-12.

4. Riddoch CJ, Bo Andersen L, Wedderkopp N et al. Physical activity levels and patterns of 9- and 15-yr-old European children. *Med Sci Sports Exerc* 2004; 36(1):86-92.

5. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Med Sci Sports Exerc* 2008;40(1):181-188.

6. Rennie KL, Wareham NJ. The validation of physical activity instruments for measuring energy expenditure: problems and pitfalls. *Public Health Nutr* 1998;1(04):265-271.

7. Trost SG. State of the art reviews: measurement of physical activity in children and adolescents. *Am J Lifestyle Med* 2007; 1(4):299-314.

8. Trost SG, Loprinzi PD. Parental influences on physical activity behavior in children and adolescents: a brief review. *Am J Lifestyle Med* 2011; 5(2):171-181.

9. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. *Med Sci Sports Exerc* 2011; 43(7):1360-1368.

10. Corder K, Ekelund U, Steele RM et al. Assessment of physical activity in youth. *J Appl Physiol* 2008; 105(3):977-987.

11. Harrell JS, McMurray RG, Baggett CD et al. Energy Costs of Physical Activities in Children and Adolescents. *Med Sci Sports Exerc* 2005; 37(2):329-336.

12. Calabro MA, Welk GJ, Eisenmann JC. Validation of the SenseWear Pro Armband algorithms in children. *Med Sci Sports Exerc* 2009; 41(9):1714-1720.

13. Arvidsson D, Slinde F, Hulthen L. Free-living energy expenditure in children using multi-sensor activity monitors. *Clin Nutr* 2009; 28(3):305-312.

14. Calabro MA, Stewart JM, Welk GJ. Validation of pattern-recognition monitors in children using doubly labeled water. *Med Sci Sports Exerc* 2013; 45(7):1313-1322.

15. Lee J-M, Kim Y, Bai Y et al. Validation of the SenseWear Mini Armband in children during semi-structure activity settings. *J Sci Med Sport* 2014; doi:10.1016/j.jsams.2014.10.004.

16. Wellek S. Introduction, Chapter 1, in *Testing statistical hypothesis of equivalence and noninferiority*, 2nd ed., Boca Raton, Fl, Chapman and Hall/CRC, 2010.

17. Batterham MJ, Van Loo CMT, Charlton KE, et al. Improved interpretation of studies comparing methods of dietary assessment: Combining equivalence testing with the limits of agreement. *Br J Nutr.* 2016. Accepted for publication.

18. Cole T, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. *Pediatr obes*. 2012; 7(4):284-294.

19. Welk GJ, McClain J, Ainsworth BE. Protocols for evaluating equivalency of accelerometry-based activity monitors. *Med Sci Sports Exerc* 2012; 44(1):S39-49.

20. Ridley K, Ainsworth BE, Olds TS. Development of a compendium of energy expenditures for youth. *Int J Behav Nutr Phys Act* 2008; 5(1):45-53.

21. Weir JBV. New methods for calculating metabolic rate with special reference to protein metabolism. *J. Physiol* 1949; 109(1-2):1-9.

22. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 327(8476):307-310.

23. Bäcklund C, Sundelin G, Larsson C. Validity of an armband measuring energy expenditure in overweight and obese children. *Med Sci Sports Exerc* 2010; 42(6):1154-1161.

**Table 1 Participants' characteristics**

| Characteristics | Mean | Median | IQR | Min - Max |
|---|---|---|---|---|
| Age (y) | 9.2 | 9.3 | 3.7 | 5.0 - 12.9 |
| Sex | | | | |
| Boys (n=28) | 49.1% | | | |
| Girls (n=29) | 50.9% | | | |
| Height (cm) | 135.9 | 137.4 | 22.2 | 104.4 - 167.0 |
| Body Mass (kg) | 32.7 | 29.3 | 16.2 | 16.6 - 56.0 |
| BMI ($kg/m^2$) | 17.1 | 16.6 | 3.3 | 14.0 - 23.8 |
| BMI percentile | 53.2 | 53.9 | 49.6 | 5.1 - 96.8 |
| Underweight (n=4) | 7.0% | | | |
| Normal weight (n=44) | 77.2% | | | |
| Overweight (n=9) | 15.8% | | | |
| Age distribution | | | | |
| 5-7 (n=19) | 33.3% | | | |
| 8-18 (n=24) | 42.1% | | | |
| 11-12 (n=14) | 24.6% | | | |
| Race | | | | |
| Caucasian (n= 54) | 94.7% | | | |
| Asian (n=3) | 5.3% | | | |

Characteristics of the participants are presented as mean ± SD, distributions of the sample are presented in percentages.

**Figure 1 revisions**

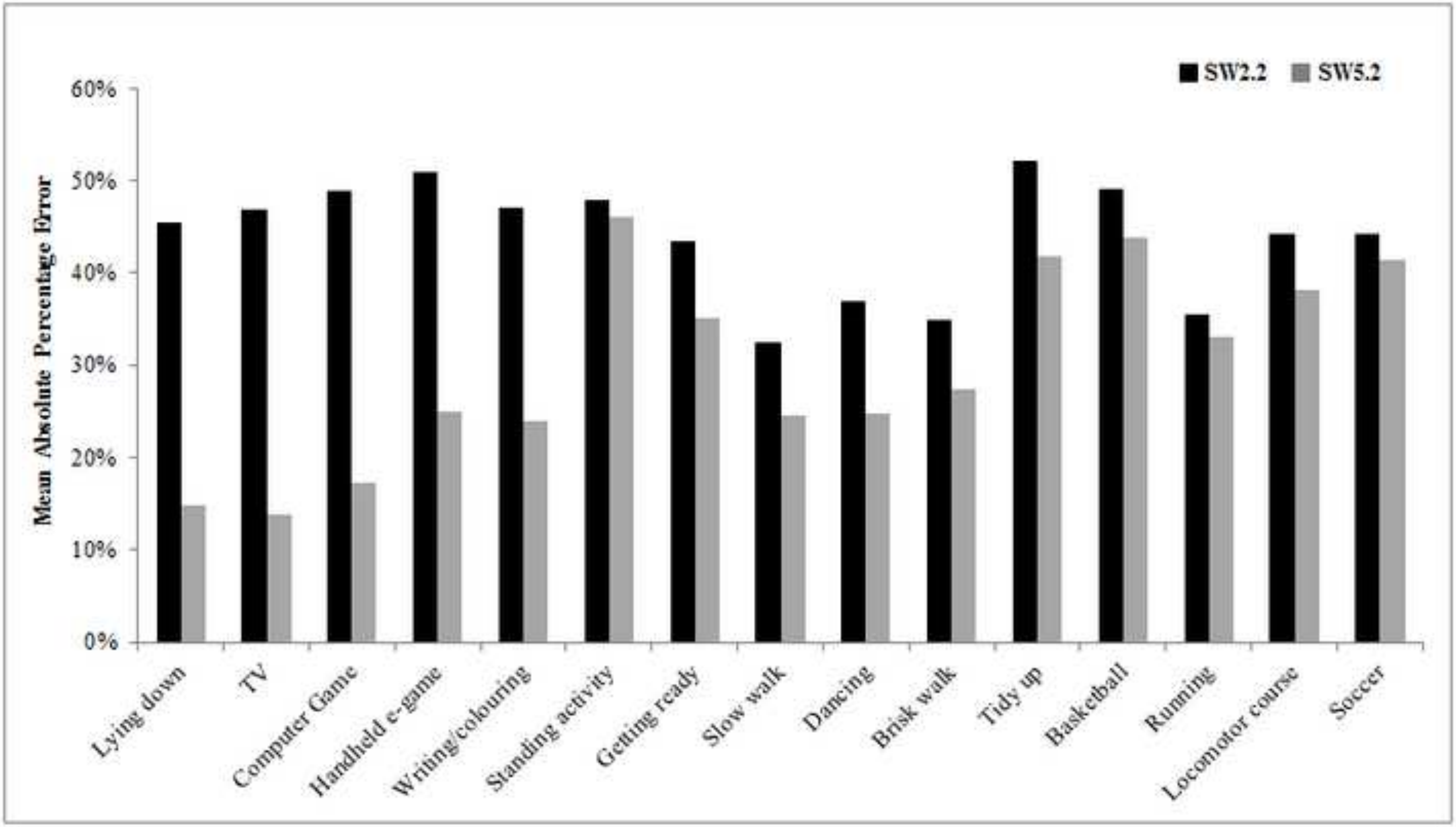
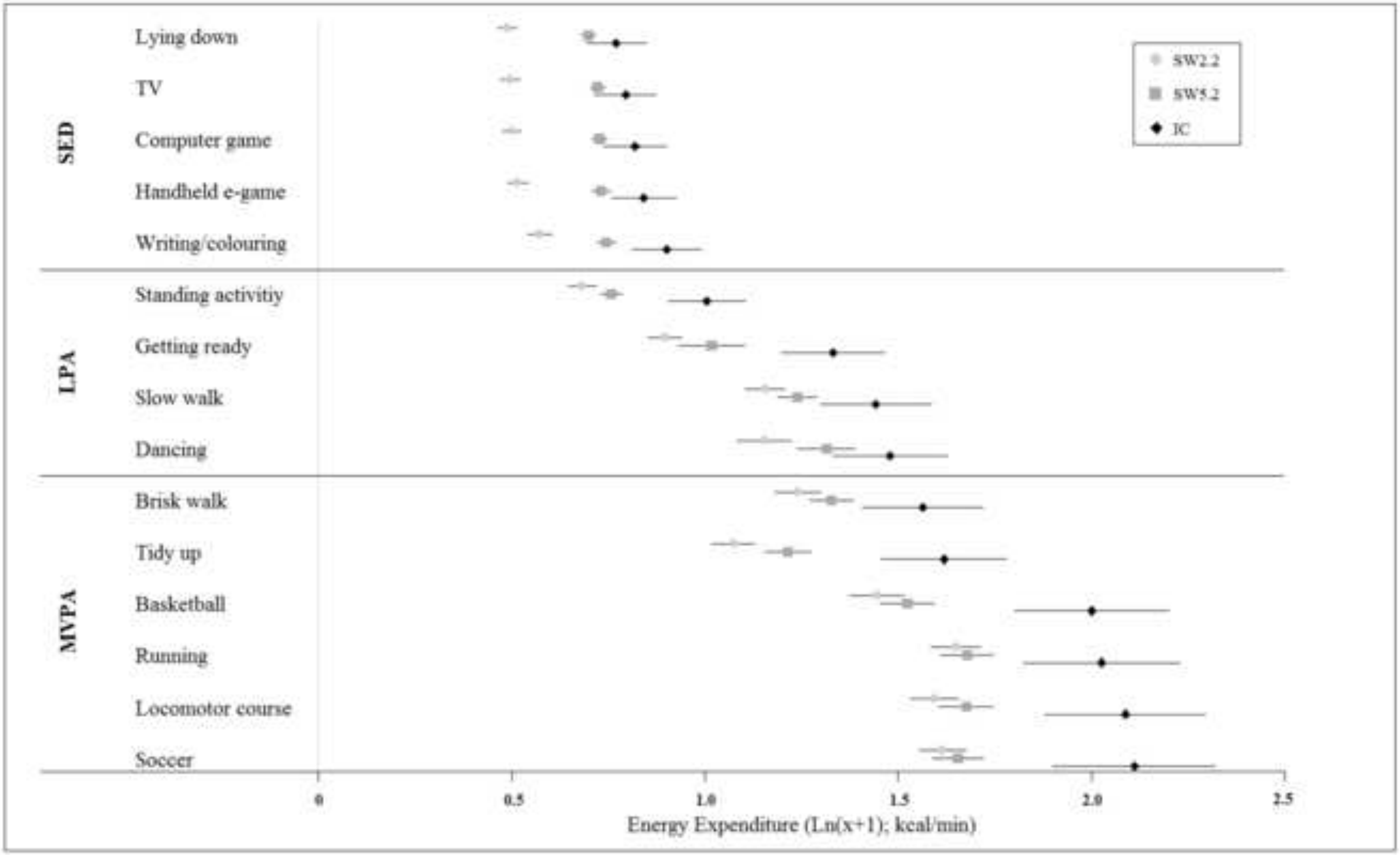


Figure 2



SED
LPA
MVPA
Lying down
TV
Computer game
Handheld e-game
Writing/colouring
Standing activitiy
Getting ready
Slow walk
Dancing
Brisk walk
Tidy up
Basketball
Running
Locomotor course
Soccer
SW2.2
SW5.2
IC
0
0.5
1.0
1.5
2.0
2.5
Energy Expenditure (Ln(x+1); kcal/min)