Validity of traditional physical activity intensity calibration methods and the feasibility of self-paced walking and running on individualised calibration of physical activity intensity in children

Eero A. Haapala1,2*, Ying Gao1,3, Anssi Vanhala1,4, Timo Rantalainen1 & Taija Finni1

There are no practical and valid methods for the assessment of individualised physical activity (PA) intensity in observational studies. Therefore, we investigated the validity of commonly used metabolic equivalent of tasks (METs) and pre-determined PA intensity classification methods against individualised PA intensity classification in 35 children 7–11-years-of-age. Then, we studied validity of mean amplitude deviation (MAD) measured by accelerometry during self-paced walking and running in assessment of individualised PA intensity. Individualised moderate PA (MPA) was defined as $\dot{V}O_2 \geq 40\%$ of $\dot{V}O_2^{\text{reserve}}$ and $\dot{V}O_2 < \text{ventilatory threshold (VT)}$ and vigorous PA (VPA) as $\dot{V}O_2 \geq VT$. We classified $> 3$–$6$ (or alternatively $> 4$–$7$) METs as MPA and $> 6$ ($> 7$) METs as VPA. Task intensities were classified according to previous calibration studies. MET-categories correctly identified 25.9–83.3% of light PA, 85.9–90.3% of MPA, and 56.7–82.2% of VPA. Task-specific categories correctly classified 53.7% of light PA, 90.6% of MPA, and 57.8% of VPA. MAD during self-paced walking discriminated MVPA from light PA (sensitivity = 67.4, specificity = 88.0) and MAD during self-paced running discriminated VPA from MPA (sensitivity = 78.8, specificity = 79.3). In conclusion, commonly used methods may misclassify PA intensity in children. MAD during self-paced running may provide a novel and practical method for determining individualised VPA intensity in children.

Free-living physical activity (PA) has been inversely associated with cardiometabolic risk in children. Moreover, higher PA intensity may confer greater cardiometabolic health benefits than lower PA intensity. Therefore, accurate assessment of PA intensity is important for informing health-related PA recommendations and in research on dose–response relationships. Nowadays, volume and intensity of PA in observational studies are typically assessed using accelerometers. However, previous studies have used fixed acceleration magnitude cut-offs to define different PA intensities without accounting for individual variation in exercise capacity. The continuing use of fixed cut-offs may have obscured our understanding on the prevalence of children accumulating recommended 60 min of daily moderate-to-vigorous PA (MVPA) and the role of PA intensity in health outcomes among youth.

1Biology of Physical Activity, Faculty of Sport and Health Sciences, University of Jyväskylä, PO Box 35, 40014 Jyväskylä, Finland. 2Physiology, Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland. 3Department of Sports Science, College of Education, Zhejiang University, Hangzhou, China. 4Department of Education, Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland. *email: eero.a.haapala@jyu.fi
Fixed acceleration intensity cut-offs based on a single absolute acceleration magnitude value have been found to underestimate PA volume and intensity in lower fit and overweight adults. Furthermore, metabolic equivalent of task (MET) is commonly used to assess PA intensity and to calibrate accelerometry cut-offs. MET approach has been criticised because METs are confounded by body composition and METs also underestimate PA intensity in overweight and obese individuals. In addition, previous calibration studies in children have provided several different cut-offs for light (LPA), moderate (MPA), and vigorous PA (VPA) leading to a large variation in the proportion of children meeting the PA recommendations. However, some evidence suggests that fixed absolute cut-offs for light (LPA), moderate (MPA), and vigorous PA (VPA) should be used.

Increasing blood lactate concentration has been considered as MPA that can be maintained prolonged periods without constantly increasing blood lactate concentration. Furthermore, PA above VT leads to an increased lactate concentration of lactate threshold, has been considered as MPA that can be maintained prolonged periods without constantly increasing blood lactate concentration. Therefore, defining PA intensity as a percentage of VO₂ reserve provides more appropriate estimate of individualised PA intensity than fixed intensity cut-offs. However, because of individual variation in ventilatory threshold (VT), metabolic responses can differ significantly between individuals who exercise at the same proportion of their VO₂ reserve. A performed below VT, a non-invasive equivalent of lactate threshold, has been considered as MPA that can be maintained prolonged periods without constantly increasing blood lactate concentration. Furthermore, PA above VT leads to an increased lactate concentration and exhaustion and is often defined as VPA. Exercise training intensity based on VT has been shown to trigger larger physiological adaptations and to reduce individual variation in physiological responses to exercise compared to exercise training intensity based on percentage of VO₂ reserve, suggesting the validity of VT in the exercise prescription and calibration of PA. Furthermore, some evidence in adults suggests that calibrating PA intensity cut-offs using lactate threshold or VT provides more accurate estimates of VPA than fixed cut-offs based on METs. However, there are no previous studies utilising VT in calibration of accelerometry cut-offs for VPA in children.

Individual calibration of cut-offs may provide superior classification accuracy for LPA, MPA, and VPA in children. Some studies have calibrated accelerometry cut-offs in a subsample of a larger study population and created fixed sample specific PA intensity cut-offs or used individually determined intensity cut-offs based on optimum walking speed and transition from walking to running during an incremental cardiopulmonary exercise test on a treadmill. However, the individualised calibration of accelerometry cut-offs against true metabolic cost or optimum walking speed and transition from walking to running during an incremental cardiopulmonary exercise test is not always feasible, because of their time consuming and resource intensive nature. On the other hand, self-selected walking speed has been found to reflect individualised MPA intensity, and we hypothesised that self-preferred walking speed may be a feasible method to individually calibrate MPA.

To improve our understanding on the role of PA intensity it is essential to investigate the validity of commonly used methods used to define PA intensity and to investigate novel methods aiming to improve PA intensity estimation in children. Therefore, we first investigated the validity of PA intensity cut-offs based on METs and task-specific calibration activities against individualised PA intensity cut-offs in children. Second, we investigated whether acceleration mean amplitude deviation (MAD) normalised for MAD measured during self-paced walking and running provide better classification accuracy for individualised PA intensity than fixed MAD cut-offs.

Results

Characteristics of participants and different physical activities. Characteristics of the participants are shown in Table 1. METs, absolute MAD, VO₂ as a % of VO₂ reserve, VO₂ as a % of VO₂ at VT, MAD relative to MAD during self-paced walking and running, and VO₂ normalised for skeletal muscle mass (SMM) or body mass (BM) increased with increasing treadmill speed and were higher during self-paced running than during self-paced walking (p < 0.001 for main effect for all comparisons, Fig. 1). MAD did not differ between walking or running on a treadmill for 6 km/h and playing hopscotch (p = 0.126). METs, VO₂ as a % of VO₂ reserve, VO₂ as a % of VO₂, and VO₂ normalised for SMM did not differ statistically significantly between walking up and down the stairs and playing hopscotch (p = 0.059 to 0.106).

Validity of PA intensity based on MET-classification and task-specific classification against individualised PA intensity classification. Compared to individualised PA intensity classification, MET-classification correctly classified 25.9% of LPA, 85.9% of MPA, and 82.2% of VPA (χ²(4) = 151, p < 0.001, Cramer’s V = 0.6, Fig. 2A,C,D). When we used > 4–7 METs to define MPA and > 7 METs to define VPA, 83.3% of LPA, 90.3% of MPA, and 56.7% of VPA (χ²(4) = 151, p < 0.001, Cramer’s V = 0.704) were correctly classified. Task-specific PA intensity classification correctly classified 53.7% of LPA, 90.6% of MPA, and 57.8% of VPA (χ²(4) = 147.3, p < 0.001, Cramer’s V = 0.595). Previously published fixed MAD cut-offs correctly classified 76.9% of LPA, 52.5% of MPA, and 54.4% of VPA (χ²(4) = 88.5, p < 0.001, Cramer’s V = 0.47, Fig. 2B).

Physical activity intensity during self-paced walking and running. The mean of VO₂ as a % of VO₂ reserve was 37.4% (min 19.3, max 54.1, SD = 8.7) and the median of VO₂ as a % of VO₂ at VT was 68.5% (min 41, max 133, IQR = 16.1) during self-paced walking (Fig. 1). Self-paced walking was categorised as LPA in 55.2%
MAD during self-paced walking and running in classifying individualised physical activity intensity. MAD as a % of MAD during self-paced walking was able to discriminate LPA from MVPA in 67.4% of the cases (Table 2). MAD as a % of MAD during self-paced running were able to discriminate VPA from MPA in 78.8% of the cases. Fixed MAD thresholds were able to discriminate LPA from MVPA in 65.8% of the cases and VPA from MPA in 66.7% of the cases. The ability to correctly identify individualised PA intensity increased when walking up and down the stairs and playing hopscotch were excluded from the data.

The ability of MAD as a % of MAD during self-paced walking and running and absolute MAD to discriminate PA intensities using different PA intensity classifications is presented in Table 2. Notably, MAD during self-paced walking and absolute MAD had higher sensitivity and specificity to correctly classify PA intensity based on task-specific calibration tasks compared to other methods used to classify PA intensity.

Discussion
We found that defining PA intensity using fixed METs and task-specific classification methods may lead to large errors in the classification of individualised PA intensity in children. We also found that MAD values measured during different laboratory activities normalised to MAD values measured during self-paced running had acceptable sensitivity to discriminate individualised VPA from MPA. Furthermore, we showed that MAD values measured during different laboratory activities normalised to MAD during self-paced running was more accurate in discriminating VPA from MPA than absolute MAD especially when more complex and intermittent physical activities, such as walking up and down the stairs and playing hopscotch, were included in the analyses.

We observed that METs, task-specific calibration, or previous MAD cut-offs [27] misclassified LPA in up to 70% and VPA in 20–45% of the cases. Furthermore, compared with 3 METs as a cut-off for MVPA and 6 METs as a cut-off for VPA, using > 4 METs and > 7 METs improved the classification accuracy to discriminate LPA from MVPA, but decreased the classification accuracy to differentiate VPA from MPA. These findings are in line with the observations in adults showing that fixed PA intensity cut-offs based on METs can lead to a significant misclassification of PA intensity and underestimate true intensity and volume of PA especially in overweight or unfit individuals [8]. To the best of our knowledge, similar findings have not been published in children. None of the previous studies utilising direct measurements of VO2 have anchored proposed PA intensity cut-offs on physiological thresholds based on VO2 data or took individual differences in VO2peak into account [13,28,29]. Furthermore, we found only small difference in MAD-based fixed PA intensity cut-offs for MPA (< 0.062 g) and VPA (0.0583) between our study and the study by Aittasalo et al. using comparable task-specific calibration method [27] suggesting that task-specific PA intensity calibration has relatively good agreement between samples, but the
Figure 1. Differences in METs, MAD, \( VO_2 \) as a % of \( VO_2 \) reserve, \( VO_2 \) as a % of \( VO_2 \) at VT, and \( VO_2 \) normalised for skeletal muscle mass (SMM) or body mass (BM) in different physical activities. 1: walking on a treadmill 4 km/h, 2: walking or running on a treadmill 6 km/h, 3: walking or running on a treadmill 8 km/h, 4: walking up and down the stairs, 5: playing hopscotch, 6: self-paced walking, 7: self-paced running. MET metabolic equivalent of task, MAD mean amplitude deviation, \( VO_2 \) oxygen uptake, VT ventilatory threshold.
cut-offs do not accurately reflect true PA intensity. These findings together indicate that common methods used to classify PA intensity may cause remarkable bias in the assessment of individualised PA intensity in children.

We observed that MAD measured during different activities normalised for MAD measured during self-paced walking or running were able to discriminate MVPA and VPA, but the sensitivity of those measures to correctly classify PA intensity was lower than in previous studies. Differences in the methods used to classify PA intensity most likely explain these differences. Most previous studies providing individualised or population specific fixed accelerometry cut-offs have used predetermined task-specific activities or evaluated the validity of the cut-offs using fixed MET-thresholds. We showed that the sensitivity and specificity of absolute MAD to discriminate task-specific PA intensity was above 80% but the ability of correctly classify individualised PA intensity was below 70%. Another reason for the discrepancy in classification accuracy with previous studies may be that our study included activities with low acceleration but high energetic demand, such as walking up and down the stairs which occurs frequently in free living, while physical activities in some previous studies have been simple ambulatory activities. Therefore, our findings suggest that fixed accelerometry cut-offs have limited validity in the assessment of individualised PA intensity.

MAD relative to MAD during self-paced running had better sensitivity to discriminate VPA from LPA and MPA compared with fixed MAD cut-offs. However, the classification accuracy of MAD during self-paced running and fixed MAD cut-offs was comparable when the data included only walking and running activities. Nevertheless, using only activities including walking and running in the calibration of PA intensity cut-offs may not reflect habitual physical activities in children and would lead to errors in the estimation of habitual PA intensity. All children, except one, operated at intensity above VT during self-paced running suggesting that anchoring...
Table 2. Receiver operating characteristics curve analyses for the accuracy of individualised and fixed cut-offs to assess moderate and vigorous physical activity in children. *Self-paced walking as independent variable b self-paced running as independent variable c The data set excluding walking up and down the stairs and playing hopscotch MVPA moderate to vigorous physical activity, VPA vigorous physical activity, AUC area under the curve, MET Metabolic equivalent of task. Within a predetermined task-specific intensity we classified walking on a treadmill for 4 km/h as LPA, running on treadmill for 6 km/h, walking up and down the stairs playing hopscotch, and walking around an indoor track on self-chosen speed as MPA, and running on a treadmill for 8 km/h and running around an indoor track on self-chosen speed as VPA according to previous calibration studies.

VPA to MAD values measured during self-paced running reflects VPA in almost all children. Fixed MAD cut-off for VPA based on our sample and from previous samples underestimated PA intensity in 33–46% of the cases while cut-offs based on MAD as a % of MAD during self-paced running misclassified 21% of the cases. Therefore, our results suggest that MAD during self-paced running provides better estimate of VPA in children than fixed MAD cut-offs.

Against our hypotheses, our results do not support the superiority of MAD during self-paced walking in the individual calibration of MVPA in children. Relatively large variation between children in VO2 during self-paced walking varying from less than 20% up to 54% of VO2 reserve may partly explain our observation. Differences in neuromuscular maturation, which may have an effect on ability to control walking intensity and walking economy, may explain this large variation in VO2 during self-paced walking in children aged 7–11 years. The strengths of the present study include a valid and simultaneous assessment of VO2 and accelerometry during different activities, assessment of true resting VO2, VO2peak and VT, and the use of individualised PA intensity in categorising LPA, MPA, and VPA. Our sample was also relatively representative to general Finnish population regarding VO2peak. We also had variable physical activities mimicking normal daily activities performed by the children. It would have been optimal to include free play and tasks where participants would have been allowed to perform free tasks in their normal environments to increase the ecological validity of the study. VO2peak and VT were assessed during a maximal cycle ergometer test and VO2peak was adjusted using the data from the treadmill running or self-paced running if higher VO2 was observed during those activities. Therefore, it is possible that we have underestimated true VO2peak in some participants. A maximal treadmill exercise test with a supramaximal validation test would have needed to measure true VO2peak to optimally reflect ambulatory activities in the present study. Furthermore, we defined MVPA as VO2 ≥ 40% of VO2 reserve. Although several previous studies have also defined MVPA as 40–55% of VO2peak, the physiological rationale for these limits to separate LPA from MVPA is lacking. Therefore, the classification accuracy of self-paced walking could have been different with another individualised intensity threshold. More research is warranted to fully understand PA intensity in children. Finally, although children were required to arrive at fasted state to the resting measurement,
they were not asked to avoid all exercise a day before the resting measurements. Therefore, exercise during the previous day could have had some minor effect on the resting VO₂ in the present study. However, children have been found to recover fast even from strenuous exercise34.

In conclusion, our results suggest that fixed MET-values or predetermined task-specific calibration activities should be used with caution in the classification of PA intensity or in the calibration of accelerometry cut-offs in children. Moreover, child-specific MAD values corresponding approximately 50% of the MAD values measured during self-paced running may provide more accurate, feasible, and practical method for individual calibration of VPA in children. Furthermore, our results suggest that fixed MAD cut-offs may be acceptable if children only walk and run, but not when they perform more complex and intermittent activities. Therefore, individualised MAD values may provide better estimates of VPA in real life setting. Further studies investigating the physiological responses to different exercise intensities and whether self-paced or optimum walking speed26, self-paced running, or the transition from walking to running can be used in individual calibration of PA intensity in large-scale observational studies in children are warranted. Finally, there is a need for further studies investigating how individualised PA volume and intensity are related to health and wellbeing in children.

Methods
Participants. This study was based on the laboratory phase of the Children’s Physical Activity Spectrum (CHIPASE) study37. A total of 35 children (21 girls, 14 boys) aged 7–11 years were recruited from local schools and volunteered to participate in the study. Children were included if they were apparently healthy and were able to perform the physical activities at moderate and vigorous intensities. Children with chronic conditions or disabilities were excluded from the study. The study protocol was approved by the Ethics Committee of the University of Jyväskylä. All children gave their assents and their parents/caregivers gave their written informed consents. The study was conducted in agreement with the Declaration of Helsinki.

Study protocol. The participants visited the laboratory three times. At the first visit, research staff explained the research protocol to children and their parents. They were also familiarised to the laboratory environment and the measurement equipment. At the second visit, children arrived at the laboratory in the morning after 10–12 h overnight fast for the assessment of anthropometrics, body composition, and resting VO₂. At the third visit, children were asked to follow six activities for 4.5 min in a random order interspersed with 1-min rest: sitting quietly, sitting while playing a mobile game, standing quietly, standing while playing a mobile game, playing hopscotch, walking up and down the stairs, and walking or running on a treadmill at 4.6, and 8 km/h. They were also asked to walk and run around an indoor track at self-chosen speed for 4.5 min. At the end of the third visit, children performed maximal cardiopulmonary exercise test on a bicycle ergometer.

Assessments. Body size and body composition. Stature was measured to the nearest 0.1 cm using a wall-mounted stadiometer. BM, SMM, fat mass, fat free mass, and body fat percent were measured by InBody 770 bioelectrical impedance device (Biospace Ltd., Seoul, Korea). Body mass index (BMI) was calculated by dividing body weight with body height squared and body mass index standard deviation score (BMI-SDS) was computed using the Finnish references38.

Oxygen uptake during rest and different physical activities. Mobile metabolic cart (Oxycon mobile, CareFusion Corp, USA) was calibrated and dead space was adjusted to 78 ml for the petite size of the face mask following the manufacturer’s recommendations. VO₂, carbon dioxide production (VCO₂) and respiratory exchange ratio (RER) were collected breath by breath and computed in non-overlapping 1 s epoch lengths. Resting VO₂ was measured in the quiet room while a child lying down as still as possible watching an age-appropriate cartoons. Resting VO₂ was determined as the mean value between the 15th and 25th minute of 30 min of supine rest when the steady state was reached39. When steady state was not observed or there were abnormal spikes in the data between 15 and 25th minute, we visually determined the steady state from the whole measurement period for further analysis. The data were visually checked among six participants and the changes made in analysis window varied between 30 s and five minutes. In physical activities, VO₂ was averaged over 2 min from the 3rd and 4th minutes of each task when the plateau in VO₂ and VCO₂ was observed39. Although the order of physical activities was randomised, we also confirmed that VO₂ returned near to baseline levels during the 1-min rest (Supplementary figure 1).

Peak oxygen uptake and oxygen uptake at ventilatory threshold. Cardiorespiratory fitness was assessed by a maximal ramp exercise test on an electromagnetically braked Ergoselect 200 K electromagnetic cycle ergometer (Ergoline, Bitz, Germany)34. The protocol included 2-min resting period sitting on an ergometer, a 3-min warm-up with a workload of 20 W, and an incremental exercise period with increase of workload either by 1 W/3 s (totalling 20 W/minute for children >150 cm), 1 W/4 s (totalling 15 W/minute for children 126–150 cm), or 1 W/6 s (totalling 10 W/minute for children ≤125 cm) until voluntary exhaustion39. The participants were asked to keep the cadence at 70–80 during the test. The test was terminated when the participant was unable to keep the cadence of 65 or required to stop. Participants were verbally encouraged to exercise until voluntary exhaustion.

Respiratory gas exchange was assessed directly by breath-by-breath method using the metabolic cart from the 2-min resting period sitting on the ergometer until the voluntary exhaustion and were averaged over 15-s periods. We defined peak cardiorespiratory capacity as the highest VO₂ achieved in the exercise test (VO₂peak) averaged over 15 s recorded during the last minute of the exercise test and normalised it for SMM. If higher VO₂ was observed during running on a treadmill for 8 km/h or during self-paced running (N = 21) than during
the maximal cycle exercise tests (N = 14), the higher VO₂ value was used as a measure of VO₂peak. Beat-by-beat heart rate (HR) was continuously recorded during the exercise test using Polar H7 HR sensor (Polar Electro, Kempele, Finland).

The cardiopulmonary exercise test was considered maximal if the primary and secondary objective and subjective criteria indicated maximal effort and maximal cardiorespiratory capacity (a plateau of VO₂ regardless of increasing workload, HR > 85% of predicted, respiratory exchange ratio > 1.00, or flushing and sweating), and the exercise physiologist supervising the exercise test considered the test maximal.

VO₂ at VT was individually determined by two exercise physiologists using modified V-slope method and any disagreements were solved by these two exercise physiologists. The VT was identified as a time point where the increase in VCO₂ was steeper than the increase in VO₂ during the maximal cardiopulmonary exercise test on a cycle ergometer. In determination of VT, we used data averaged over 15 seconds. VO₂ at VT was verified utilising the equivalents for V̇O₂/VCO₂ and V̇E/V̇CO₂. According to equivalent method VO₂ at VT was defined as a rate of VO₂ where V̇O₂/VCO₂ begins to increase without an increase in V̇E/V̇CO₂.

**Accelerometry during different physical activities.** Movement was measured by triaxial accelerometer (X6-1a, Gulf Coast Data Concepts Inc., Waveland, USA). We used raw acceleration data in actual g-units with the high range up to 6 g with 16-bit A/D conversion, and sampling at 40 Hz. The resultant acceleration of the triaxial accelerometer signal was calculated from \(\sqrt{x^2 + y^2 + z^2}\), where x, y and z were the measurement sample of the raw acceleration signal in x-, y-, and z-directions. The X6-1a accelerometer has been shown to produce congruent results with the ActiGraph GT3X accelerometer in children. The MAD was calculated from the resultant acceleration in non-overlapping 1 s epoch. MAD describes the mean distance of data points around the mean (\(\frac{1}{n} \sum_{i=1}^{n} |r_i - \bar{r}|\), where n was the number of samples in the epoch, \(\bar{r}\) is the ith resultant sample within the epoch and \(r\) is the mean resultant value of the epoch). The mean of the 1 s MAD values (g) were calculated in the exercise physiologist supervising the exercise test considered the test maximal.

**Definitions of intensity in different physical activities.** Individualised intensity classification. We defined individualised PA intensity using VO₂ reserve and VO₂ at VT. VO₂ reserve was calculated as \((\text{VO₂}_{\text{during PA task}} - \text{VO₂}_{\text{during rest}}) \times 100\). LPA was defined as VO₂ below 40% and MPA as VO₂ ≥ 40% of VO₂ reserve to < VO₂ at VT. VPA was defined as VO₂ at or above VO₂ at VT. MVPA was defined as activity VO₂ ≥ 40% of VO₂ reserve.

**MET-based intensity classification.** MET values were computed as VO₂ measured during the physical activities/VO₂ during supine rest. PA intensity during different tasks was categorised as follows: LPA was defined as > 1.5–3 METs, MPA as > 3–6 METs, and VPA as > 6 METs. We also performed additional analyses using the cut-offs of > 4 METs for MPA and VPA. The absolute child-specific MAD values representing the cut-offs for MPA and VPA, the absolute child-specific MAD value measured during self-paced walking or self-paced running should be used and the appropriate multiplier provided in the Table 2 should be applied.

**Predetermined task-specific intensity classification.** We classified walking on a treadmill for 4 km/h as LPA, running on treadmill for 6 km/h, walking up and down the stairs playing hopscotch, and walking around an indoor track on self-chosen speed as MPA, and running on a treadmill for 8 km/h and running around an indoor track on self-chosen speed as VPA according to previous calibration studies. We then classified as tasks excluding tasks considered LPA. We also investigated the agreement of previously published fixed MAD cut-offs for LPA (< 332 mg), MPA (332 mg), and VPA (558.3 mg) based on predetermined task-specific intensity classification provided by Aittasalo et al. with individualised PA intensity.

**Statistical methods.** Basic characteristics between girls and boys were compared using Student’s t-test for normally distributed continuous variables and Mann–Whitney U-test for skewed continuous variables. We investigated differences in METs, absolute MAD, VO₂ as a % of VO₂ reserve, VO₂ as a % of VO₂ at VT, MAD relative to MAD during self-paced walking and running, and VO₂ normalised for SMM or BM in different physical activities using mixed-effects repeated measures ANOVA. We investigated the validity of MET-based PA intensity classification, task-specific intensity classification, and previously published fixed MAD cut-offs against individualised PA intensity cut-offs using \(\chi^2\) test and Cramer’s V.

We used receiver operating characteristics (ROC) curves to investigate the optimal cut-off for MAD as a % of MAD during self-paced walking or running and absolute MAD to differentiate LPA from MPA and MPA from VPA using individualised PA intensity classification. The area under the curve (AUC) is considered a measure of the effectiveness of the predictor variable to correctly discriminate MVPA from LPA and VPA from MPA (sensitivity) and to correctly discriminate LPA from MVPA and MPA from VPA (specificity). An AUC of 1 represents the ability to perfectly identify MVPA or VPA from other intensities, whereas an AUC of 0.5 indicates no greater predictive ability than chance alone. The optimal cutoff was determined by the Youden index, which is the maximum value of 1 that is computed as: sensitivity + specificity - 1.

Student˘s t-test, the Mann–Whitney U-test, and the \(\chi^2\) test were performed using the SPSS Statistics, Version 23.0 (IBM Corp., Armonk, NY, USA). The data were visualised and the mixed-effects repeated measures ANOVA were performed by the GraphPad Prism, version 8.0.2 (GraphPad Software, Inc., San Diego, CA, USA). The ROC curve analyses were performed using MedCalc Statistical Software, Version 16.1 (MedCalc Software bvba, Ostend, Belgium).
Data availability
The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 11 December 2019; Accepted: 10 June 2020
Published online: 03 July 2020

References
1. Poitras, V. J. et al. Systematic review of the relationships between objectively measured physical activity and health indicators in school-aged children and youth. Appl. Physiol. Nutr. Metab. 41, 197–239 (2016).
2. Vaïstö, J. et al. Longitudinal associations of physical activity and sedentary time with cardiometabolic risk factors in children. Scand. J. Med. Sci. Sports 29, 113–123 (2019).
3. Collings, P. J. et al. Cross-sectional associations of objectively-measured physical activity and sedentary time with body composition and cardiorespiratory fitness in mid-childhood: the PANIC Study. Sport. Med. 47, 769–780 (2017).
4. Moore, J. B. et al. Associations of vigorous-intensity physical activity with biomarkers in youth. Med. Sci. Sports Exerc. 49, 1366–1374 (2017).
5. Corder, K., Ekelund, U., Steele, R. M., Wareham, N. J. & Brage, S. Assessment of physical activity in youth. J. Appl. Physiol. 105, 977–987 (2008).
6. Stratton, S. J. et al. Guide to the assessment of physical activity: Clinical and research applications: a scientific statement from the American Heart Association. Circulation 128, 2259–2279 (2013).
7. Migueles, J. H. et al. Comparability of published cut-points for the assessment of physical activity: Implications for data harmonization. Scand. J. Med. Sci. Sports 29, 566–574 (2019).
8. Kujala, U. M. et al. Physical activity: absolute intensity versus relative-to-fitness-level volumes. Med. Sci. Sports Exerc. 49, 474–481 (2017).
9. Tompuri, T. T. Metabolic equivalents of tasks are confounded by adiposity, which disturbs objective measurement of physical activity. Front. Physiol. 6, 1–6 (2015).
10. Vähä-Yppä, H. et al. Validation of cut-points for evaluating the intensity of physical activity with accelerometry-based Mean Amplitude Deviation (MAD). PLoS ONE 10, 1–13 (2015).
11. Ravagnani, F. et al. Application of the rosetta stone to understand how much MVPA preschoolers accumulate: A systematic review. J. Sci. Med. Sport 20, 849–855 (2017).
12. Gao, Y. et al. Children’s physical activity and sedentary time compared using assessments of accelerometer counts and muscle activity level. PeerJ 6, e5437 (2018).
13. Evenson, K. R., Catellier, D. J., Gill, K., Ondrak, K. S. & McMurray, R. G. Calibration of two objective measures of physical activity for children. J. Sports Sci. 26, 1557–1565 (2008).
14. Fredson, P., Pober, D. & Jänz, K. F. Calibration of accelerometer output for children. Med. Sci. Sports Exerc. 37, S523–S530 (2005).
15. Trost, S. G., Loprinzi, P. D., Moore, R. & Pfeiffer, K. A. Comparison of accelerometer cut points for predicting activity intensity in youth. Med. Sci. Sport Exerc. 43, 1360–1368 (2011).
16. Garber, C. E. et al. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med. Sci. Sports Exerc. 43, 1334–1359 (2011).
17. Meyer, T., Gabriel, H. W. H. & Kindermann, W. Is determination of exercise intensities as percentages of VO2max or HRmax adequate?. Med. Sci. Sport. Exerc. 31, 1342–1345 (1999).
18. Baldwin, J., Snow, R. J. & Febbraio, M. A. Effect of training status and relative exercise intensity on physiological responses in men. J. Sci. Med. Sport 20, 849–855 (2017).
19. Pate, R. R., Pfeiffer, K. A., Dowda, M., McIver, K. L. & Almeida, M. J. Validation and calibration of an accelerometer in preschool children. Prev. Med. 42, 566–574 (2006).
20. Trost, S. G., Pate, R. R., Pfeiffer, K. A., Sirard, J. R. & Dowda, M. Calibration and evaluation of an objective measure of physical activity intensity levels in postmenopausal women. Eur. J. Sport Sci. 19, 539–548 (2019).
21. Murtagh, E. M., Boreham, C. A. G. & Murphy, M. H. Speed and exercise intensity of recreational walkers. Eur. J. Appl. Physiol. 97, 997–1004 (2002).
22. Puyau, M., Adolph, A., Vohra, F. E. & Butte, N. Validation and calibration of physical activity monitors in children. Obes. Res. 10, 150–157 (2002).
23. Pate, R. R., Pfeiffer, K. A., Dowda, M., McIver, K. L. & Almeida, M. J. Validation and calibration of an accelerometer in preschool children. Obesity 14, 2009, 2000–2001 (2006).
24. Van Cauwenberghe, E., Labarque, V., Trost, S. G., De Bourdeaudhuij, I. & Cardon, G. Calibration and comparison of accelerometer cut points in preschool children. Int. J. Pediatr. Obes. 6, e582–e589 (2011).
25. Frost, S. G., Pate, R. R., Pfeiffer, K. A., Sirard, J. R. & Dowda, M. Calibration and validation of an objective measure of physical activity in preschool children. J. Phys. Act. Heal. 2, 345–357 (2015).
26. Radnor, J. M. et al. The influence of growth and maturation on stretch-shortening cycle function in youth. Sport. Med. 48, 57–71 (2018).
27. Luoto, N. & et al. Cardiorespiratory fitness, respiratory function and hemodynamic responses to maximal cycle ergometer exercise in girls and boys aged 9–11 years: the PANIC Study. Eur. J. Appl. Physiol. 115, 235–243 (2015).
28. Bilodeau, S. et al. Metabolic and fatigue profiles are comparable between prepubertal children and well-trained adult endurance athletes. Front. Physiol. 9, 387 (2018).
29. Gao, Y. et al. Sedentary thresholds for accelerometry-based mean amplitude deviation and electromyography amplitude in 7–11 years old children. Front. Physiol. 10, 997 (2019).
36. Saari, A. et al. New Finnish growth references for children and adolescents aged 0 to 20 years: Length/height-for-age, weight-for-length/height, and body mass index-for-age. *Ann. Med.* **43**, 235–248 (2011).
37. Ventham, J. C. & Reilly, J. J. Reproducibility of resting metabolic rate measurement in children. *Br J Nutr* **81**, 435–437 (1999).
38. Saint-Maurice, P. E., Kim, Y., Welk, G. J. & Gaesser, G. A. Kids are not little adults: what MET threshold captures sedentary behavior in children? *Eur. J. Appl. Physiol.* **116**, 29–38 (2016).
39. Godfrey, S., Davies, C. T. M., Wozniak, E. & Barnes, C. A. Cardio-respiratory response in normal children. *Clin. Sci.* **40**, 419–431 (1971).
40. Machado, F. A. & Denadai, B. S. Validity of maximum heart rate prediction equations for children and adolescents. *Arq Bras Cardiol.* **97**, 136–140 (2011).
41. Ruby, B. C. et al. Validity and reliability of combining three methods to determine ventilatory threshold. *Med. Sci. Sports Exerc.* **33**, 1841–1848 (2003).
42. Laukkanen, A., Pesola, A., Havu, M., Sääkslahti, A. & Finni, T. Relationship between habitual physical activity and gross motor skills is multifaceted in 5- to 8-year-old children. *Scand. J. Med. Sci. Sports* **24**, 102–110 (2013).
43. Vähä-Ypää, H., Vasankari, T., Husu, P., Suni, J. & Sievänen, H. A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. *Clin. Physiol. Funct. Imaging* **35**, 64–70 (2015).
44. Perkins, N. J. & Schisterman, E. F. The inconsistency of ‘optimal’ cutpoints obtained using two criteria based on the receiver operating characteristic curve. *Am. J. Epidemiol.* **163**, 670–675 (2006).

**Acknowledgements**

We thank all CHIPASE Study participants for their time and dedication to the study and the Ms Tanja Niemi, MSc and Mr Martti Melin, MSc for helping carrying out the assessments. This study was funded by Ministry of Education and Culture of Finland (OKM/59/626/2016). The authors declare no conflicts of interest.

**Author contributions**

E.A.H. analysed the data and drafted the manuscript. Y.G., A.V., and T.R. collected and processed the data for analyses. T.F. planned the research and received funding and contributed to planning the manuscript and interpreting the results and reviewed the manuscript. All authors provided significant intellectual contribution to the manuscript.

**Competing interests**

The authors declare no competing interests.

**Additional information**

Supplementary information is available for this paper at [https://doi.org/10.1038/s41598-020-67983-7](https://doi.org/10.1038/s41598-020-67983-7).

**Correspondence** and requests for materials should be addressed to E.A.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit [http://creativecommons.org/licenses/by/4.0/](http://creativecommons.org/licenses/by/4.0/).

© The Author(s) 2020