COMMENTARY   **Open Access**

Population-level physical activity surveillance in young people: are accelerometer-based measures ready for prime time?

Stewart G Trost

With the promotion of physical activity in young people, an established global health priority [1], it is critically important for governments and public health agencies to have a clear understanding of the proportion of children and adolescents meeting guidelines for physical activity and sedentary behaviour. Historically, the assessment of physical activity in large scale population-level surveillance systems has been limited to self-report measures. However, among youth, self-report methods are subject to significant social desirability and recall bias [2, 3]. Younger children, in particular, have difficulty recalling their past behaviour accurately; and struggle to understand the concepts of physical activity frequency, intensity, duration, and type [4]. Proxy self-reports completed by parents or caregivers are one solution, but this method is also subject to recall bias since respondents can only report on the time they are in contact with the child [2, 3]. In light of the limitations of self-report methods, device based physical activity measures such as accelerometers have become the preferred method in studies involving children and young people [2, 5].

Despite the ubiquitous use of accelerometer-based motion sensors in physical activity studies involving children and adolescents, the application of accelerometers in population-level physical activity surveillance systems has been seriously questioned [6]. Citing methodological limitations related to questionable validity, between-monitor variability, multiple sets of conflicting intensity-based cut-points, and bias resulting from monitoring non-compliance, Pedisic and Bauman [6] concluded that without appropriate standardisation protocols, the adoption of accelerometers for large-scale physical activity surveillance was premature. They further concluded that

accelerometer-based physical activity measures should not substitute, but only supplement self-report information systems; and that self-report methods should be the primary assessment tool in physical activity surveillance systems.

The newly published paper by Steene-Johannessen et al. [7] goes a long way to dispel Pedisic and Bauman's notion that accelerometer-based measures are not yet ready for "prime time" when it comes to population-level physical activity surveillance. Based on the reanalysis of accelerometer data from more than 47,000 young people from 30 different studies across 18 European countries, the authors present what could be argued as the most detailed and comprehensive descriptions of physical activity and sedentary behaviour levels in European children and adolescents to date. The results indicate that approximately two-thirds of European children and adolescents do not meet the daily 60 min moderate-to-vigorous physical activity (MVPA) recommendation, with youth from southern European countries being at significantly greater risk for inactivity than those residing in central or northern Europe. The findings are concerning to say the least and underscore the need for effective policies and programs to promote physical activity and reduce sedentary behaviour in children and young people.

Through the implementation of a carefully conceived accelerometer data processing protocol, Steene-Johannessen and colleagues [7] mitigate many of the concerns identified in the Pedisic and Bauman review [6]. The analysis was delimited to studies using an ActiGraph accelerometer on the hip, epoch length was standardised to 60 s, data from midnight to 6:00 am was excluded to minimise any bias associated with between-study variation in monitoring protocols and minimal wear time criteria, and point estimates of MVPA and sedentary behaviour were based on a single set of cut-point values, regardless of child age. Nevertheless,

Correspondence: s.trost@qut.edu.au
Institute of Health and Biomedical Innovation at QLD Centre for Children's Health Research, Queensland University of Technology, Level 6, 62 Graham Street, South Brisbane, QLD 4101, Australia

the numerous study limitations described in the discussion section indicate that not all the methodological limitations identified in the Pedisic and Bauman review were addressed adequately.

A significant limitation of the Steene-Johannessen study is the reliance on proprietary count-based metrics - making it impossible to integrate data from studies deploying different brands of accelerometers. The solution to this problem is to discontinue the practice of calculating and reporting count-based metrics and applying the currently available methods and metrics based on the raw acceleration signal [8–10]. It is acknowledged that Steene-Johannessen included data from studies that were conducted long before raw accelerometer data from the ActiGraph could be readily accessed, limiting their ability to use this approach. However, if future accelerometer data pooling projects are to provide legitimate between-study and cross-country comparability, the more than two decade old practice of applying cut-points to processed activity counts must be phased out.

In the Pedisic and Bauman review [6], the existence of multiple sets of conflicting intensity-based cut-points was identified as a major methodological weakness limiting the application of accelerometers in population-level surveillance studies. Steene-Johannessen and colleagues partially address this issue by applying, in a standardised manner, an intensity-based cut-point with established evidence of validity in school-aged youth [11, 12]. However, there is growing recognition that the relationship between accelerometer counts and energy expenditure is highly dependent on the activities included the calibration study; and that cut-points derived from a single regression model or Receiver Operating Characteristic curve cannot adequately characterise physical activity intensity across a wide range of physical activities [13]. In an independent evaluation of ActiGraph cut-points for youth, the Evenson thresholds were found to have the least physical activity intensity classification error of all the cut-points tested [12]. However, it is important to note that the Evenson cut-points still misclassified MVPA as light-intensity physical activity 20% of the time, and that light intensity physical activities were misclassified as sedentary at least 40% of the time [12]. Moreover, given that the relationship between activity counts and energy expenditure in children under five differs substantially to that observed in adolescent youth [14], the application of the Evenson cut-points in children aged 2- to 5-years by Steene-Johannessen must be questioned.

Over the last decade, there has been a shift from count-based thresholds to machine learning activity classification and energy expenditure estimation algorithms based on features extracted from raw accelerometer signals [15]. When applied to youth, machine learning approaches have shown to provide more accurate predictions of physical activity intensity [13, 16]. Moreover, in contrast to cut-point methods, which only estimate time spend in MVPA, physical activity classification models can predict time spent in specific activity types (e.g., walking, running, dancing, cycling) or broader activity classes (e.g., active games or sports) [13, 16]. This enables researchers in the public health and exercise sciences to explore a greater variety of physical activity metrics as well as examine age-related differences in movement behaviours that are not confounded by developmental differences in the relationship between accelerometer counts and energy expenditure.

To date, the uptake of machine learning methods by public health researchers has been slow, primarily because of the need to collect and process large quantities of raw accelerometer signal using specialised software; and partly because of concerns that machine learning models trained on laboratory-based activities trials do not generalise well to free living scenarios [17]. As machine learning accelerometer data processing methods evolve and the required computer platforms enabling public health researchers to apply machine learning methods become available, it is anticipated that future physical activity surveillance studies will address the aforementioned limitations of cut-point methods by implementing potentially more accurate and versatile machine learning accelerometer data processing methods.

In their review, Pedisic and Bauman [6] identify between-study variations in monitoring protocols and non-compliance (non-wear) as key methodological issues limiting the use of accelerometer in physical activity surveillance studies. In their study, Steene-Johannessen and colleagues [7] partially address this limitation by excluding all accelerometer data recorded from midnight to 6:00 am and excluding monitoring days with less than eight hours of valid wear time. While these decision rules are not without precedent, it is likely that at a significant percentage of participants were awake and/or wearing the monitor during the excluded time periods. Furthermore, as noted by the authors, the inclusion of studies implementing 24 h monitoring protocols may have led to an overestimation of sedentary time, given the longer daily wear periods and the misclassification of sleep as sedentary time. It can be argued that the issue of non-wear is a legacy issue from objective monitoring studies requiring participants to wear an accelerometer on the hip during the waking hours. Because hip mounted accelerometers are typically placed on snug fitting elastic belts and worn over clothing, non-compliance and insufficient wear time were frequent problems in these studies. To improve compliance and minimise missing data due to non-wear, more and more studies are implementing continuous 24-h monitoring protocols with wrist mounted

accelerometers. Wrist mounted accelerometers are easier to wear for extended periods and allow investigators to evaluate compliance with more contemporary 24-h movement guidelines which require concurrent monitoring of physical activity, sedentary behaviour, and sleep [18]. With wrist mounted accelerometers, it is critically important that researchers compute physical activity metrics using more sophisticated accelerometer data processing methods [16, 19]. Simple cut-point approaches applied to either processed activity counts or raw acceleration signal (i.e., Euclidean Norm Minus One - ENMO) provide misleading estimates of movement behaviour because they do not account for upper limb movements during sedentary or stationary light-intensity activities [20, 21]. In investigations where sitting time is of primary interest, assessments of posture with thigh mounted accelerometers, alone or in combination with other placements, should be considered [22, 23].

The findings of the Steene-Johannessen study highlight the long-standing methodological issue of how to operationalise compliance with physical activity guidelines in accelerometer-based studies [24]. Steene-Johannessen and colleagues [7] operationalise meeting guidelines as accumulating an average of > 60 min of MVPA across all valid monitoring days; but other studies have applied more stringent criteria by requiring participants to accumulate ≥60 min on every monitoring day [25]. It can be argued that, based on the wording of the guidelines, children are required to reach the target of 60 min of MVPA on each and every day. If this is the case, then significantly more than two-thirds of European children and adolescents are insufficiently active for health benefit. Alternatives to the "average over all days" and "all days" methods are "the most days" methods (e.g., meeting the 60 min recommendation on > 50% of monitoring days) and the "Child x Day" method used by Cooper at colleagues [25] in which the percentage of valid monitoring days with ≥60 min of MVPA is calculated and interpreted as the probability that a randomly selected child on a randomly selected day met the guideline [24]. It is acknowledged that all methods have advantages and disadvantages. However, the difficulty associated with comparing prevalence estimates from studies applying different approaches highlights the need for consensus on how physical activity guidelines for young people should be operationalised.

The article by Steene-Johannessen and colleagues [7] highlights the utility of accelerometer-based measures in physical activity surveillance studies involving children and adolescents. The findings are significant and identify the promotion of physical activity as a priority concern for public health authorities across Europe. Yet, it is important to acknowledge that the accelerometer data processing methods applied in this study have been implemented with only minor modifications for more than two decades, despite significant advances in wearable sensor technology and artificial intelligence over the same time period. If accelerometer-based measures of physical activity and sedentary behaviour are to be accepted as best practice methodology in large scale population-level surveillance studies, public health researchers must be willing to adopt more contemporary monitoring protocols and apply new accelerometer data processing methods.

References
1. World Health Organization. Global Recommendations on Physical Activity for Health. Geneva: World Health Organisation; 2010.
2. Trost SG. Measurement of physical activity in children and adolescents. Am J Lifestyle Med. 2007;1:299–314.
3. Hildebrand M, Ekelund U. The assessment of physical activity. In: Armstrong N, van Mechelen W, editors. Oxford Textbook of Children's Sport and Exercise Medicine. 3rd ed. New York: Oxford University Press; 2017. p. 303–13.
4. Trost SG, Morgan AM, Saunders R, Felton G, Ward DS, Pate RR. Children's understanding of the concept of physical activity. Pediatr Exerc Sci. 2000;12: 293–9.
5. Cliff DP, Reilly JJ, Okely AD. Methodological considerations in using accelerometers to assess habitual physical activity in children aged 0-5 years. J Sci Med Sport. 2009;12:557–67.
6. Pedisic Z, Bauman A. Accelerometer-based measures in physical activity surveillance: current practices and issues. Br J Sports Med. 2015;49:219–23.
7. Steene-Johannessen J, et al. Variations in accelerometry measured physical activity and sedentary time across Europe – harmonized analyses of 47,498 children and adolescents. IJBNPA. (in press).
8. Wijndaele K, Westgate K, Stephens SK, et al. Utilization and harmonization of adult accelerometer data: review and expert consensus. Med Sci Sports Exerc. 2015;47:2129–39.
9. Hildebrand M, van Hees V, Hansen BH, Ekelund U. Age group comparability of raw accelerometer output from wrist- and hip-worn monitors. Med Sci Sports Exerc. 2014;46:1816–24.
10. Hildebrand M, Hansen BH, van Hees V, Ekelund U. Evaluation of raw acceleration sedentary thresholds in children and adults. Scand J Med Sci Sports. 2017;27:1814–23.
11. Evenson KR, Cattellier D, Gill K, Ondrak K, McMurray RG. Calibration of two objective measures of physical activity for children. J Sports Sci. 2008;26: 1557–65.
12. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut-points for predicting activity intensity in youth. Med Sci Sports Exerc. 2011;43:1360–8.
13. Trost SG, Wong WK, Pfeiffer KA, Zheng Y. Artificial neural networks to predict activity type and energy expenditure in youth. Med Sci Sports Exerc. 2012;44(9):1801–9.

14. Trost SG, Drovandi CC, Pfeiffer K. Developmental trends in the energy cost of physical activities performed by youth. J Phys Act Health. 2016; 13(6 Suppl 1):S35.
15. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med. 2014;48:1019–23.
16. Trost SG, Cliff D, Ahmadi M, Van Tuc N, Hagenbuchner M. Sensor-enabled activity class recognition in preschoolers: hip versus wrist data. Med Sci Sports Exerc. 2018;50(3):634–41.
17. Ahmadi MN, Brookes D, Chowdhury A, Pavey T, Trost SG. Free-living evaluation of laboratory-based activity classifiers in preschoolers. Med Sci Sports Exerc. 2019. https://doi.org/10.1249/MSS.0000000000002221.
18. Okely AD, Ghersi D, Hesketh KD, et al. A collaborative approach to adopting/adapting guidelines - The Australian 24-Hour Movement Guidelines for the early years (Birth to 5 years): an integration of physical activity, sedentary behavior, and sleep. BMC Public Health. 2017;20;17(Suppl 5):869.
19. Trost SG, Zheng Y, Wong WK. Machine learning for activity recognition: hip versus wrist data. Physiol Meas. 2014;35(11):2183–9.
20. VAN Loo CMT, Okely AD, Batterham MJ, Hinkley T, Ekelund U, Brage S, Reilly JJ, Trost SG, Jones RA, Janssen X, Cliff DP. Wrist acceleration cut points for moderate-to-vigorous physical activity in youth. Med Sci Sports Exerc. 2018; 50(3):609–16.
21. Trost SG, Rice K, Pfeiffer KA. Comparison of wrist accelerometer cut-points for classifying physical activity intensity in youth. J Sci Med Sport. 2017; 20(suppl1):e104–5.
22. Crowley P, Skotte J, Stamatakis E, Hamer M, Aadahl M, Stevens ML, Rangul V, Mork PJ, Holtermann A. Comparison of physical behavior estimates from three different thigh-worn accelerometers brands: a proof-of-concept for the prospective physical activity, sitting, and sleep consortium (ProPASS). Int J Behav Nutr Phys Act. 2019;16(1):65.
23. Stewart T, Narayanan A, Hedayatrad L, Neville J, Mackay L, Duncan S. A dual-accelerometer system for classifying physical activity in children and adults. Med Sci Sports Exerc. 2018;50(12):2595–260.
24. Olds T, Ridley K, Wake M, Hesketh K, Waters E, Patton G, Williams J. How should activity guidelines for young people be operationalised? IJBNPA. 2007;4:43.
25. Cooper AR, Goodman A, Page AS, et al. Objectively measured physical activity and sedentary time in youth: the International children's accelerometry database (ICAD). IJBNPA. 2015;12:113.

Publisher's Note