Published in final edited form as:

Am J Prev Med. 2017 June ; 52(6): 836–838. doi:10.1016/j.amepre.2017.03.001.

Understanding and Interpreting Error in Physical Activity Data: Insights from the FLASHE Study

Gregory J. Welk, PhD¹, Pedro F. Saint-Maurice, PhD², Youngwon Kim, PhD³, Laura D. Ellingson, PhD¹, Paul Hibbing, MS⁴, Dana L. Wolff-Hughes, Ph.D.⁵, and Frank M. Perna, EdD, PhD⁵

¹Department of Kinesiology, Iowa State University, Ames, Iowa ²Metabolic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland

³MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, United Kingdom ⁴Department of Kinesiology, University of Tennessee, Knoxville, Tennessee

⁵Health Behaviors Research Branch, Behavioral Research Program, National Cancer Institute, Bethesda, Maryland

The Family Life, Activity, Sun, Health, and Eating (FLASHE) Study involved the collection of data on an array of lifestyle behaviors,¹ but considerable efforts were made to obtain the best available data on physical activity (PA) and sedentary behavior (SB) because both have unique and independent contributions to health. FLASHE used accepted “best practices” by incorporating both monitor- and report-based methods (e.g., self-report measures) to enable triangulation of outcomes.^{2,3} Monitor-based methods provide a sound way to objectively compare overall amount of PA and SB in large population studies as the improved precision increases statistical power.^{4,5} Stronger associations with health indicators have generally been observed with monitor-based data^{6–9} but report-based measures still provide valuable information about context and individual perceptions of behavior.³ Considerable research has been conducted to refine and improve assessments of PA and SB in youth with both methods, but less attention has been placed on how to interpret the data and how to link the outcomes. This brief commentary provides an overview of the PA assessment approaches used in the FLASHE Study as well as recommendations for interpreting the integrated data from the various PA indicators. The challenges with processing and interpreting data from monitor-based devices have been well documented.¹⁰ For example, it is well established that estimates vary widely depending on the type of monitor, the approaches used to screen the data, the sampling frequency, and the choice of algorithm or prediction equation. Consistent with many large surveys (e.g., National Health and Nutrition Examination Survey and United Kingdom Biobank study), and overall trends in accelerometry research, FLASHE used wrist-worn devices to improve compliance with the monitoring protocol and collected

Address correspondence to: Gregory J. Welk, PhD, Department of Kinesiology, 257 Forker Building, 534 Wallace Ave., Iowa State University, Ames IA. gwelk@iastate.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

raw accelerometer data instead of using processed “activity counts”. These features add new layers of complexity to evaluation of monitor-based research^{11,12} but there are clear advantages with respect to standardization and data harmonization.¹³ A challenge with this new form of data is that little is currently known about the best ways to process and interpret raw data from wrist-worn devices. Existing methods and algorithms developed for hip-worn monitors do not apply to these data, so it is important to understand the variability in outcomes generated with different processing methods to further advance work in this area.

The paper by Kim et al.¹⁴ provides novel insights on these issues by directly evaluating the variability in outcomes from the leading methods currently available for wrist-worn Actigraph data. As expected, considerable variability was observed in estimates of PA from the various methods but, regardless of method, no meaningful differences in PA levels were observed between boys and girls. The results could be due to the sample, the uniqueness of wrist-worn data, or the available processing methods, so additional research is clearly needed in this area. The lack of consensus in data processing standards make it difficult to draw absolute conclusions about behavior, but the objective indicators of PA and SB provide useful outcomes for examining associations with other health or behavior indicators. The linked data from parent–teen dyads also offer new opportunities to understand family influences on lifestyle behaviors. Troiano and colleagues¹² recently recommended caution with the use of accelerometry-based monitors for surveillance applications, so emphasis should be on the evaluation of patterns in the data and relevant associations with health or behavioral correlates.

Another consideration with interpreting PA and SB data in studies like FLASHE is in reconciling the commonly observed discrepancies between estimates from monitor- and report-based data.¹⁵ Discrepancies in the estimates reported from the two measures should be somewhat expected because monitor-based methods record absolute or total movement whereas report-based measures reflect a person’s perception of the movement and the relative intensity.³ A novel aspect of the FLASHE Study is that it incorporated a calibrated report-based tool called the Youth Activity Profile (YAP) that was developed specifically to produce estimates of PA and SB that correspond closely with group-level estimates obtained from monitor-based data.¹⁵

The paper by Saint-Maurice et al.¹⁷ describes the specific steps and procedures used to calibrate the YAP within the FLASHE Study. The calibration of the reported PA behavior enables more uniform reporting of outcomes in large studies like FLASHE because the estimates are equated to the monitor values through the calibration. However, it is important to note that the estimates are directly dependent on the assumptions and methods used in the calibration process. In FLASHE, the estimates of PA and SB were designed to match estimates obtained using the algorithm of Crouter and colleagues¹⁸ as analyzed by Kim et al.¹⁴ in this issue. The estimates from this approach had better agreement with other laboratory and field-based data in several evaluation substudies conducted by our research team on behalf of National Cancer Institute prior to the processing of the FLASHE data (reports available upon request). However, because of the calibration, the estimates of PA and SB from the YAP must be interpreted in the same way as other monitor-based data and not be directly used to infer compliance with PA guidelines. The advantage of the calibration

procedure is that it converts the report-based data into a more useable outcome measure that can be used in a diverse range of subsequent analyses using the full sample of youth. For example, the paper in this supplement by D'Angelo and colleagues¹⁹ utilized the calibrated YAP estimates to evaluate psychosocial correlates of PA behavior within the FLASHE sample. The FLASHE data include the processed YAP estimates of PA and SB to enable additional research with other indicators and variables in FLASHE. Processed monitor-based data will also be available to enable comparisons with alternative methods and metrics.

The papers by Kim et al.¹⁴ and Saint-Maurice and colleagues¹⁷ provide useful summaries of the PA and SB estimates obtained from the FLASHE youth sample. They also highlight key measurement considerations that need to be addressed to further refine assessment methods. This is important because careful attention to the strengths and limitations of available methods is necessary to advance research on PA and SB. A key point is that with monitor-based methods it is currently only possible to “assess” and not “measure” PA and SB. These tools are designed to capture PA and SB constructs based on the degree of movement and there is considerable error in these estimations, regardless of approach. The need for precision and the implications of measurement error certainly depend on the way the data are used and the research questions being addressed. For example, the implications of systematic and random error depend on whether PA and SB data are used as exposure or outcome variables.²⁰ In general, with current methods, it is best to focus on the underlying patterns in the data (and associations with other health-related behaviors and outcomes) rather than on the absolute estimates of PA and SB. Going forward, greater attention is needed to incorporate measurement error models that can quantify and control for error in both monitor- and report-based methods. As assessment methods and processing strategies advance, it will be possible to assess PA and SB with greater accuracy and precision, which will then help to answer more-complex public health and epidemiology questions.

Acknowledgments

Pedro Saint-Maurice was a postdoctoral researcher with Iowa State University during the completion of the Family Life, Activity, Sun, Health, and Eating project and during the preparation of the manuscript.

No financial disclosures were reported by the authors of this paper.

References

1. Masse LC, Lytle LA. Advancing knowledge of parent-child dyadic relationships about multiple cancer preventive health behaviors: The National Cancer Institute Family Life, Activity, Sun, Health, and Eating (FLASHE) study. *Am J Prev Med*. 2017 In press.
2. Haskell WL. Physical activity by self-report: a brief history and future issues. *J Phys Act Health*. 2012; 9(Suppl 1):S5–10. <https://doi.org/10.1123/jpah.9.s1.s5>. [PubMed: 22287448]
3. Troiano RP, Pettee Gabriel KK, Welk GJ, et al. Reported physical activity and sedentary behavior: why do you ask? *J Phys Act Health*. 2012; 9(Suppl 1):S68–75. <https://doi.org/10.1123/jpah.9.s1.s68>. [PubMed: 22287450]
4. Strath SJ, Kaminsky LA, Ainsworth BE, et al. Guide to the assessment of physical activity: Clinical and research applications: a scientific statement from the American Heart Association. *Circulation*. 2013; 128(20):2259–2279. <https://doi.org/10.1161/01.cir.0000435708.67487.da>. [PubMed: 24126387]

5. Arem H, Keadle SK, Matthews CE. Invited Commentary: Meta-Physical Activity and the Search for the Truth. *Am J Epidemiol*. 2015; 181(9):656–658. <https://doi.org/10.1093/aje/kwu472>. [PubMed: 25855648]
6. Celis-Morales CA, Perez-Bravo F, Ibanez L, et al. Objective vs. self-reported physical activity and sedentary time: effects of measurement method on relationships with risk biomarkers. *Plos One*. 2012; 7(5):e36345. <https://doi.org/10.1371/journal.pone.0036345>. [PubMed: 22590532]
7. Atienza AA, Moser RP, Perna F, et al. Self-reported and objectively measured activity related to biomarkers using NHANES. *Med Sci Sport Exer*. 2011; 43(5):815–821. <https://doi.org/10.1249/MSS.0b013e3181fdcf32>.
8. Hearst MO, Sirard JR, Lytle LA, et al. Comparison of three measures of physical activity and associations with blood pressure, HDL and body composition in a sample of adolescents. *J Phys Act Health*. 2012; 9(1):78. <https://doi.org/10.1123/jpah.9.1.78>. [PubMed: 22232509]
9. Belcher BR, Moser RP, Dodd KW, et al. Self-Reported Versus Accelerometer-Measured Physical Activity and Biomarkers among NHANES Youth. *J Phys Act Health*. 2015; 12(5):708–716. <https://doi.org/10.1123/jpah.2013-0193>. [PubMed: 24909801]
10. Matthews CE, Hagstromer M, Pober DM, et al. Best Practices for Using Physical Activity Monitors in Population-Based Research. *Med Sci Sport Exer*. 2012; 44(suppl 1):S68–S76. <https://doi.org/10.1249/MSS.0b013e3182399e5b>.
11. Staudenmayer J, Zhu WM, Catellier DJ. Statistical Considerations in the Analysis of Accelerometry-Based Activity Monitor Data. *Med Sci Sport Exer*. 2012; 44(suppl 1):S61–S67. <https://doi.org/10.1249/MSS.0b013e3182399e0f>.
12. Troiano RP, McClain JJ, Brychta RJ, et al. Evolution of accelerometer methods for physical activity research. *Br J Sports Med*. 2014; 48(13):1019–1023. <https://doi.org/10.1136/bjsports-2014-093546>. [PubMed: 24782483]
13. Wijndaele K, Westgate K, Stephens SK, et al. Utilization and Harmonization of Adult Accelerometry Data: Review and Expert Consensus. *Med Sci Sport Exer*. 2015; 47(10):2129–2139. <https://doi.org/10.1249/MSS.0000000000000661>.
14. Kim Y, Hibbing P, Saint-Maurice PF, Ellingson LD, et al. Surveillance of youth physical activity and sedentary behavior with wrist accelerometry. *Am J Prev Med*. 2017 In press.
15. Tucker JM, Welk GJ, Beyler NK. Physical activity in U.S.: adults compliance with the Physical Activity Guidelines for Americans. *Am J Prev Med*. 2011; 40(4):454–461. <https://doi.org/10.1016/j.amepre.2010.12.016>. [PubMed: 21406280]
16. Saint-Maurice PF, Welk GJ. Validity and Calibration of the Youth Activity Profile. *Plos One*. 2015; 10(12) <https://doi.org/10.1371/journal.pone.0143949>.
17. Saint-Maurice PF, Kim Y, Hibbing P, Oh A, Perna FM, Welk GJ. Calibration and validation of the Youth Activity Profile: The FLASHE study. *Am J Prev Med*. 2017 In press.
18. Crouter SE, Flynn JJ, Bassett DR. Estimating Physical Activity in Youth Using a Wrist Accelerometer. *Med Sci Sport Exer*. 2015; 47(5):944–951. <https://doi.org/10.1249/MSS.0000000000000502>.
19. D'Angelo H, Fowler SL, Nebeling LC, et al. Adolescent physical activity: Moderation of individual factors by neighborhood environment. *Am J Prev Med*. 2017 In press.
20. Hutcheon JA, Chiolerio A, Hanley JA. Random measurement error and regression dilution bias. *BMJ*. 2010; 340:c2289. <https://doi.org/10.1136/bmj.c2289>. [PubMed: 20573762]