Strategies for Dealing with Missing Accelerometer Data

Samantha Stephens, PhD[a,*], Joseph Beyene, PhD[b], Mark S. Tremblay, PhD[c], Guy Faulkner, PhD[d], Eleanor Pullnayegum, PhD[e], Brian M. Feldman, MD, MSc, FRCPC[f]

KEYWORDS

• Imputation • Missing • Data • Physical activity • Measurement

KEY POINTS

- Missing data poses a threat to the validity, reliability, and generalizability of data from physical activity trials.
- Consideration for the type and amount of missing data is necessary to select appropriate imputation methods.
- Multiple imputation should be used to replace missing physical activity data because it has been shown to give the most unbiased estimates.

INTRODUCTION

Participation in recommended levels of physical activity has been associated with important clinical outcomes in youth and adults with rheumatologic conditions, such as improved physical function, joint pain, disability, and overall quality of life.[1–4] Yet, youth and adults with rheumatologic conditions are highly inactive.[1–3,5] It is necessary to understand the optimal dose of physical activity to understand and measure its effect on health in youth and adults with rheumatologic conditions. Accelerometry, an

Disclosures: The authors have no conflicts of interest or disclosures to declare in relation to this article.

[a] Neurosciences and Mental Health, Pediatric M.S., Neuroinflammatory Disorders Program, Center for Brain and Mental Health, Peter Gilgan Centre for Research and Learning, The Hospital for Sick Children, 686 Bay Street, Room 8.9830, Toronto, Ontario M5G 0A4, Canada; [b] Department of Clinical Epidemiology and Biostatistics, 208 Michael G. DeGroote Centre for Learning, McMaster University, 1280 Main Street W., Hamilton, Ontario L8S 4K1, Canada; [c] Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Ontario, Canada; [d] Canadian Institutes of Health Research, Public Health Agency of Canada, Applied Public Health, University of British Columbia, D. H. Copp Building, Room 4606 2146 Health Sciences Mall, Vancouver, BC V6T1Z3, Canada; [e] Child Health Evaluative Sciences, The Hospital for Sick Children, Public Health Sciences, The University of Toronto, Toronto, Ontario, Canada; [f] Child Health Evaluative Sciences, The Hospital for Sick Children, Department of Pediatrics, Institute of Health Policy Management and Evaluation, The Dalla Lana School of Public Health, The University of Toronto, Toronto, Ontario, Canada
* Corresponding Author.
E-mail address: Samantha.stephens@sickkids.ca

Rheum Dis Clin N Am 44 (2018) 317–326
https://doi.org/10.1016/j.rdc.2018.01.012
rheumatic.theclinics.com

objective and unobtrusive method for determining physical activity patterns, has proven a valid outcome measurement tool that overcomes many of the methodologic limitations of subjective methods (eg, self-reports), such as recall bias.[6,7] Despite quality control methods a particular challenge to physical activity measurement with accelerometers is missing data.

Missing data is a universal measurement problem that threatens the integrity of study results by calling into question the interpretation, reliability, and generalizability of the findings. Regardless of the outcome measurement tool used, missing data can introduce bias and challenge the interpretation of study findings. Accelerometer missing data poses a unique problem, however, because many times partial data exist for the participant involved in the study. For example, data may be missing for a position of the day (eg, a participant took off the monitor for a brief period) or entire days of data may be missing (eg, 2 out of 7 days).

To properly address missing data it is important to differentiate between the mechanisms of missingness that one may encounter, because imputation strategies to deal with missingness are dependent on the mechanism. According to Little and Rubin[8] there are three mechanisms of missing data: (1) missing completely at random (MCAR), (2) missing at random (MAR), and (3) not missing at random (NMAR). Data considered MAR suggest that the pattern of missingness is systematically related to some unobserved characteristic of the missing variable.[8] To classify missing data as MCAR, it must be established that the missingness is completely unrelated to the variables that are being studied.[8] Thus, the proportion of the total sample with missing data cannot be differentiated from the sample with complete data. Finally, NMAR is designated when the missingness is neither MAR nor MCAR, or the reason for the missing observations is related to the unobserved outcome.

In this article we describe the problem of missing data in the context of conducting research involving the measurement of physical activity via accelerometry as an outcome. We describe the different statistical approaches that have been used, and describe the benefits and detriments of each of the statistical approaches with consideration given to the missingness mechanism.

PHYSICAL ACTIVITY OUTCOME MEASUREMENT AND DATA PROCESSING

New technologies have been developed to more accurately measure physical activity at the individual and population level in an objective manner. The accelerometer, one such technology, has gained popularity as an objective outcome measure of physical activity that allows the user to derive time stamped and sequenced data on body movements in real time while the subject is in their own environment. Many studies use accelerometers to capture and describe physical activity behaviors, from quantifying time spent in varying intensities of physical activity to describing patterns of physical activity, such as weekdays versus weekends in varying populations. Accelerometry has proven to be a valid and reliable measure of physical activity measurement in healthy children and adults.[7,9–11] Quality control and data reduction procedures have been developed to enhance robustness and validity of data captured from the accelerometer.[12] For example, decision rules, algorithms, and methods to identify nonwear periods and spurious data have been established and some have been incorporated into software packages.[13,14]

Strategies proposed to process data from accelerometers including wear time algorithms and valid days analysis introduce bias, and limit the generalizability of the findings and comparability between studies.[14] In a recent narrative review with commentary on the methodology of prospective observational studies determining the relationship