Reviewer's report

Title: Estimates of adherence and error analysis of physical activity data collected via accelerometry in a large study of free-living adults

Version: 1 Date: 6 November 2007

Reviewer: Mark Tremblay

Reviewer's report:

General

This paper explored estimates of adherence and error analyses of physical activity data collected by accelerometry in free-living adults. The authors describe the adherence in their study population and examine methods to assess the inherent error associated with nonwear and how imputation procedures can be used to reduce this error. The paper is certainly of value for those using accelerometers to assess free living physical activity levels in adults. The paper raises some important issues, though most are not unique. The paper relies heavily on the Catellier et al. (2005) paper, and for the most part appears to use an adult population to examine the same general issues. These issues were also addressed by Esliger et al. (Journal of Physical Activity and Health 3:366-383, 2005). The paper provides some addition to the literature in this area however but with some additional work it's contribution could be significantly enhanced.

--------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. line 66-67 - the National Health and Nutrition Examination Survey in the U.S. have adherence rates available through their publicly available dataset on several thousand Americans, representative of the general population in the U.S. Insufficient information is provided by the authors to get a sense of the representativeness of the population used and the likelihood of the findings being generalizable to other adult populations with different ethics mixes, ages, climate situations, social norms, etc. In this regard, some comparison to a representative dataset, or at least some discussion on this issue would help inform readers of the potential utility of the recommendations provided in the manuscript. Please provide some information on demographics of the sample (socioeconomic status/mix; country of data collection).

2. line 99-101 - 20 consecutive zeros may or may not be appropriate - how sensitive were the findings to this assumption? What if this rule was 30 consecutive zeros? 40? If the results are very sensitive to this, the more important factor becomes this decision rule, not the type or use of imputation. Some assessment of this would be helpful.
3. Similar to the point above, a program was created to predict wake and sleep time because the authors do not trust the diaries. These values differ significantly from the sleep time recorded by the participants (41 min). How sensitive are the subsequent procedures to this program? What if it were modified to more closely approach what respondents recorded? What if you accepted the recored times of the respondents? Would these changes strengthen or weaken the effects of imputation?

4. Are the results similar between men and women? Between weekdays and weekend days? Again, trying to get at the generalizability or universality of the recommendations made by the authors. Greater efforts to test how robust their recommendations are should be made.

5. The choice to use an exponential decay model is fine, but needs to be tested against other models (when you assess only one, how can you recommend it? it was only compared to doing nothing). Also, physical activity behaviour is only somewhat autocorrelated with neighboring times (for example most work outs start and end rather abruptly). The first figure in figure 2 shows thennoise associated with most movement behaviours. Other models should be attempted to justify the choice made.


7. Many if not most studies using accelerometers determine the number of minutes of moderate and vigorous PA as their primary outcome variable. What happens to these results when the recommendations are applied? This would be important to assess and would likely be affected by some of the recommendations and the reduced variation resulting from the recommended data imputation procedures.

8. How were saturation values dealt with (counts of 32767)? These spurious values can have a significant effect on average count values (obviously).

9. In the absence of comparsion analyses the paper cannot justify recommending the kriging (or any other) procedure (line 258). This whole paragraph is based on weak support.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. Please provide research hypotheses.

2. What time of year was the data collection - discuss how this may affect findings.

3. Were incentives provided to the participants (e.g. honorarium for wearing the AM)? Discuss how this may impact results and applicability of findings.

4. Please descibe how the accelerometers were calibrated before use in teh
study.

5. line 120 - citing reference 28 this way reads like these authors developed SAS, which I don’t think they did - they can probably clarify by saying "developed a computer program using SAS (spell out) ..."

6. Simulation D should be added as another figure in figure 2.

7. line 178: It is stated that TEE was calculated using the equations by Weber et al. (2001). It would be more appropriate to reference the original papers (cited by Weber et al. in their methods) which describe both the equation used to calculate energy expenditure from CO2 production (Weir, 1949) and the sampling protocol used: multi-point method (Schoeller 1983, 1988). Also, the manufacturer details of the isotope ratio mass spectrometer used should be provided in brackets (page 9, line 177).

8. Please provide a reference for Akaike's Information Criterion.

9. line 266 - define "profound understanding of imputational statistics"

10. lines 280-281 - where are the results to support this?

11. line 435 - adherence of what?

12. line 440 - add "averages of each individual" before "hours"

13. Figure 1 - are these averages per person over the days worn by each person - or are these based on each day of each person? If it is not the latter, it should be. It should be by person/day - this needs to be made clear in the paper.

14. Table 1 - I do not understand the "# of simulations/ subject" column - what does this mean/refer to?

Discretionary Revisions (which the author can choose to ignore)

1. line 88 change "read" to "store"

2. line 89 put "(epoch)" after "time"

3. line 92 13-15 continuous(?) days?

4. Please define the REE methods a little more. What metabolic cart was used? Minor points: Any control for physical activity in previous 24 hours? Any control for smoking or alcohol intake?

5. line 205 include "after" before "omitting"

What next?: Unable to decide on acceptance or rejection until the authors have
responded to the major compulsory revisions

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests.