Assessing PA with self-report: A methodological overview

Louise C. Mâsse, PhD
University of British Columbia

Presented at
Measurement of Active and Sedentary Behaviors: Closing the Gaps in Self-Report Methods
National Institutes of Health, Bethesda, MD • July 21, 2010
Overview of talk

• Utility of self-report
• Type of studies PA self-report are used in
• Type evidence needed in these studies
• How we are currently assessing the validity of PA self-report
• Discuss potential issues to consider in the validation process
Utility of Self-Report?

• Advantages
  – Most practical & economical
  – Only feasible method in large scale studies
  – Best at measuring context & type of PA performed

• Limitations
  – Utility in some context debated
  – Validity & reliability less than objective methods

[Sallis & Saelens (2000); Shephard (2003); Westerterp (2009)]
TYPE OF STUDIES PA SELF-REPORT ARE USED IN
Type of Studies

- Epidemiological studies
- Surveillance studies
- Intervention studies
# Type of Studies

- **Epidemiological studies**
  - Assess relations between PA and health

- **Surveillance studies**

- **Intervention studies**
Type of Studies

Epidemiological studies

- Assess relations between PA and health

Surveillance studies

- Monitor levels of PA and patterns of change at the population level

Intervention studies

Type of Studies

- **Epidemiological studies**
  - Assess relations between PA and health

- **Surveillance studies**
  - Monitor levels of PA and patterns of change at the population level

- **Intervention studies**
  - Generally aim to detect small changes in PA
Methodological Consideration

- Validation sample
- Study design
- Validation criterion
- Statistical procedures
Methodological Consideration

- Validation sample
- Study design
- Validation criterion
- Statistical procedures

Content validation process - completed
- Qualitative interviews
- Cognitive interviews
- Pilot testing....
Level of evidence

Association

Agreement

Sensitivity to change

Criterion
Level of evidence & interpretation

• Association
  – Indicate whether different people who responded to the questions did more or less PA (Ranking of the scores on the PA self-report ).
Level of evidence & interpretation

• **Association**
  – Indicate whether different people who responded to the questions did more or less PA (Ranking of the scores on the PA self-report).

• **Agreement**
  – Allow the scores on the PA self-report to be interpreted in terms of actual amount of PA the person does. Actual min would correspond to the behavior.
Level of evidence & interpretation

• Association
  • Indicate whether different people who responded to the questions did more or less PA (Ranking of the scores on the PA self-report).

• Agreement
  – Allow the scores on the PA questionnaire to be interpreted in terms of actual amount of PA the person does. Actual min would correspond to behavior
<table>
<thead>
<tr>
<th>Type of studies</th>
<th>Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epidemiological</td>
<td>Association</td>
</tr>
<tr>
<td>Surveillance</td>
<td>Relative change over time</td>
</tr>
<tr>
<td></td>
<td>Association</td>
</tr>
<tr>
<td></td>
<td>Actual amount of PA - % meet recommendations</td>
</tr>
<tr>
<td></td>
<td>Agreement</td>
</tr>
<tr>
<td>Intervention</td>
<td>Sensitivity to change</td>
</tr>
<tr>
<td>Type of studies</td>
<td>Level of evidence</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Epidemiological</td>
<td>Association</td>
</tr>
<tr>
<td>Surveillance</td>
<td>Association</td>
</tr>
<tr>
<td>Relative change over time</td>
<td>Association</td>
</tr>
<tr>
<td>Actual amount of PA - % meet recommendations</td>
<td>Agreement</td>
</tr>
<tr>
<td>Intervention</td>
<td>Sensitivity to change</td>
</tr>
</tbody>
</table>

Group level validity

Individual level validity
WHAT ARE WE CURRENTLY DOING?
ASSOCIATION
• Relationship between PA and health exist

Level of PA

- High
- Moderate
- Low

CVD risk factors

- Low
- Moderate
- High

Requires valid relative ranking for PA self-report
Associations – Epidemiological studies

• Appropriate methods to assess ranking?

• Limited in their ability to account for measurement errors in the validation process
Associations – Epidemiological studies

• Statistical models that account for both measurement errors (Buonaccorsi, 2000) and correct for attenuations in epidemiological studies are not new but deserve more attention.

Mixed models
Family of statistical procedures
Associations — Epidemiological studies

Full sample N = 1000
Uncorrected association = .25

PA
Self-report

CVD
risk factors
Associations – Epidemiological studies

Full sample N = 1000
Uncorrected association = 0.25
Corrected association = 0.35

Correction for Attenuation
N = 200

PA
Self-report

CVD
risk factors

Criterion
(subset of sample)
AGREEMENT
Surveillance Studies

- Relative ranking – Association
  - If reliability is acceptable then it can be used to assess group level change from year to year
  - Assume no secular trend affecting patterns and that the questions remain the same

Can detect group level change
Cannot assess % meeting recommendation
Surveillance Studies

• Absolute interpretation – Agreement
  – % meeting PA recommendation need actual min

• Issue to consider
  – Do we have a valid criterion?
Possible Criterion

• Accelerometers
• VO2 max
• Doubly Labeled water
• Pedometers
• 6 min walk test
• Run test
• Heart rate
• Body fat
• FEV....
Useful Criterion

• Accelerometers
• VO2 max
• Doubly Labeled water
• Pedometers
• 6 min walk test
• Run test
• Heart rate
• Body fat
• FEV....
Criterion

- Accelerometry
  - Min above a certain count can be compared
  - Does not capture all activities (walking based)
• Doubly Labeled water
  – TEE = RMR + TEF + AEE (Kcal metric)
  – Self-report PAEE = min * MET value * BW/60
  – High intensity activities account for 25% of AEE need to assess moderate and inactivity (Westerterp, 2009)
Criterion

• Doubly Labeled water
  – TEE = RMR + TEF + AEE (Kcal metric)
  – Self-report PAEE = min * MET value * BW/60
  – High intensity activities account for 25% of AEE
    need to assess moderate and inactivity
• Transforming the instruments on the same metric does not mean that both instruments (self-report and criterion) are capturing the same thing.
Agreement – Surveillance studies

- Appropriate methods to assess agreement?
- Other methods to quantify agreement?

PA Self-report

Criterion

Bland & Altman, group comparison (t-test – Anova – MANOVA), kappa, Kendall tau-b
Quantifying Agreement

- Intra-class correlation

\[ ICC - \text{consistency} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2} \]

\[ ICC - \text{agreement} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_m^2 + \sigma_e^2} \]

- Receiver Operating Characteristic (ROC) – (application Trost et al, 2006)

% meeting PA rec.

Criterion

Self-report
Quantifying Agreement

- Intra-class correlation

\[
ICC - \text{consistency} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_e^2}
\]

- Receiver Operating Characteristic (ROC) – (application Trost et al, 2006)

\[
ICC - \text{agreement} = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_m^2 + \sigma_e^2}
\]

% meeting PA rec.

Criterion

Self-report
SHOULD WE FOCUS ON QUANTIFYING BIAS (OVER REPORTING) AS A WAY OF IMPROVING AGREEMENT?
Correcting for bias — systematic error

• Developing prediction equations
  – PA min from criterion
  – PA categories from criterion (Meeting / not meeting rec.)

• Challenges
  – Modeling the type of errors we have (Plankey 1995 BMI)
  – Drift phenomenon - verifying that bias is consistent over time or sample (Gorber 2010 BMI)

• Learning from the educational field

• Measurement error models
SENSITIVITY TO CHANGE
(INDIVIDUAL LEVEL VALIDITY)
Intervention studies

• Is the measure able to detect differences pre / post interventions?

  Change in PA
  Self-report

  Change Criterion

  Anova, t-test, Wilcoxon test

• Appropriate methods to assess change?

• Other methods to quantify change?
## Quantifying change

### Issues with this approach

<table>
<thead>
<tr>
<th>Pre</th>
<th>Post</th>
<th>Comparing Pre and Post scores</th>
<th>Issues with this approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Criterion (VO₂ max) p&lt;.05</td>
<td>Those who changed on the criterion might not be the same who changed on the PA self-report.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PA self-report p&lt;.05</td>
<td></td>
</tr>
</tbody>
</table>
## Quantifying change

<table>
<thead>
<tr>
<th>Pre</th>
<th>Post</th>
<th>Comparing Pre and Post scores</th>
<th>Issues with this approach</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Criterion (VO₂ max) p&lt;.05</td>
<td>Those who changed on the criterion might not be the same who changed on the PA questionnaire.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PA self-report p&lt;.05</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pre</th>
<th>Post</th>
<th>Pre and Post change scores</th>
<th>Alternative methods</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Criterion (VO₂ max)</td>
<td>Change score can be compared statistically using either “correlational methods or agreement methods.” Agreement only possible if the metrics are comparable.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PA self-report</td>
<td></td>
</tr>
</tbody>
</table>
Quantifying change

• Validating change scores with “correlational methods” (regression methods)
  – Accurately rank the change with PA self-report

• Validating change scores with “agreement methods”
  – Change score can be meaningfully quantified (in min for example)
UTILITY OF ADVANCED PSYCHOMETRIC METHODS
• Two key assumptions are problematic
  – Local independence
  – Essential unidimensionality
Assumptions

PA domains

- Household
- Yard work
- Family
- Volunteer
- Transportation
- Exercise / Sports
- Employment
- Miscellaneous

• 1-dominant dimension that explains about 40% of the variance. This assumes high correlations among domains.

• Errors are not correlated after you account for dominant dimension.
Assumptions

PA domains

- Household
- Yard
- Family
- Community
- Transportation
- Exercise / Sports
- Employment / unpaid work
- Miscellaneous

Correlation likely to be moderate

Correlation likely to be low or non-existent
Assumption test – real data

- Minority women (N=250)
  - Essential unidimensionality test
    - Factor analysis forcing a 1-factor solution
Assumption test – real data

- Correlations among domains
  - Highest correlation = 0.207
  - Next highest correlation = 0.137

- 1-factor explain 16% of total variance
  - A 2\textsuperscript{nd} factor would explain 14% of total variance
  - Not all items load on 1\textsuperscript{st} dimension

Clear violation of the assumptions
CAN ADVANCED PSYCHOMETRIC METHODS BE USED TO VALIDATE PA MEASURES?
Can we estimate the % meet who PA recommendation?
Take home

• Before you select a PA instrument for your study – make sure it has been validated for how you intend to interpret the data.
What evidence do we have?

<table>
<thead>
<tr>
<th>Category</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Associations</td>
<td>• Lots at this level</td>
</tr>
<tr>
<td>Agreement</td>
<td>• None</td>
</tr>
<tr>
<td>Sensitivity to change</td>
<td>• Some but weaker design</td>
</tr>
</tbody>
</table>
WILL BETTER STATISTICAL METHODS IMPROVE OUR SELF-REPORT INSTRUMENTS?