Presented at the International Society for Pharmacoeconomics and Outcomes Research Annual Meeting • May 7 – 10, 2023 • Boston, MA USA • In-person

CO182

Email DanielSerrano@openhealthgroup.com for more information

Nikolai D. Ayasse[1]; Frederick B. Barnes[1]; Daniel Serrano[1]

[1]OPEN Health, Bethesda, MD, USA

Click or scan the QR code or use the URL for more information https://publications.openhealthgroup.com/behind-the-science-ispor-us-2023/

Developing Methods to Validate Outcome Measures Captured Via Wearable Activity Trackers

OPEN HEALTH

Background

- Wearable activity trackers (WATs) are increasingly used for passive mobility data collection
- The data emerging WATs is non-normal
- Percentage Activity Time (PAT) metrics are biased because of uncontrolled wear time

Objective

- **To develop classical test theory (CTT) estimates of Test-retest reliability (TRR) and construct validity (CV) that are unbiased in the PAT metric**
- **Specifically, the two-way random intraclass correlation coefficient ICC(2,1) and the Pearson correlation coefficient, respectively**

Methods

1. To address known bias in covariance parameters for PAT metric, the event/trial binomial (ETB) model was employed:
   - The ETB models numerator and denominator outcome variables and computes the correctly weighted ratio from these variables
   - In doing so, it avoids equating ratios of 0.5 derived from different numerator/denominator combinations

Methods (Cont'd)

- For example:
  - For 50 minutes of moderate activity out of 6 hours (360 min) of WAT wear-time, the outcome is: 50 / 360
  - For 6 minutes of moderate activity out of 43.2 min of WAT wear-time, the outcome is 6/43.2
  - Both yield a PAT of 13.9%; however, one PAT has more information than the other – ignoring this biases variance
2. However, the ETB yields estimates in generalized metric, necessitating conversion to correlation-based metrics

ETB ICC(2,1) / ICC(A,1)

- $ICC(2,1) = (MS_b - MS_e) / (MS_b + (k\text{-}1)*MS_e + 2*(MS_j - MS_e) / n)$
- $MS_e = \sqrt{3} / \pi$
- $MS_b = k*MS_{Subj} + MS_e$
- $MS_j = n* MS_{Time} + MS_e$
- k = number of timepoints
- n = number of subjects
- MS_{Subj} = Variance Component for subject
- MS_{Time} = Variance Component for time

Methods (Cont'd)

Convert ETB OR to Pearson's r

- **Express Cohen's D as a function of the ETB OR and the standardized logistic error variance**
  - $D = OR * (\sqrt{3} / \pi)$
  - $r = D / \sqrt{(D^2 + 4)}$

3. Test accuracy of conversions with simulations, by simulating:
   - 1000 draws of size n=500, each composed of two 7-day time-series
   - Compliance and wear-time range: 1-7 days and 1-1440 minutes
   - Outcome: minutes of "vigorous" activity
   - Concurrent validator & expected Pearson correlation with outcome of r = 0.40
   - PGIS anchor variable; within "stable" group ICC(2,1) = 0.80

Results

Table 1 | Estimation Accuracy for ICC(2,1)

Convergence	Avg. ICC(2,1)	% Bias	MSE	Avg. N "no change"
No, 2.8%	.	.	.	153.7
Yes, 97.2%	0.788	-1.5%	0.000	154.4

Results (Cont'd)

Table 2 | Estimation Accuracy for Cross-sectional ETB-based Pearson r for Construct Validity

Convergence	Avg. Pearson r	% Bias	MSE
No, 0.0%	.	.	.
Yes, 100%	0.39	-2.5%	0.000

Conclusions

- **Some COA data are, by definition, generalizations of distributions routinely employed**
  - **Percent Activity Time weighted by the minutes of wear time is one such example**
- **It is possible to adapt CTT frameworks for validation to generalized distributions**
  - **Herein we extended estimation of the ICC(2,1) and Pearson's Correlation to ETB data**
  - **Bias for all estimates across 1000 replications was uniformly below 10%**
  - **Both estimators underestimated the true value slightly: maximum of 2.5% underestimation**
- **Longitudinal ETB models can fail to converge (2.8%, n=28, in this simulation failed)**
  - **Analysts should not interpret results from models that failed to converge**
- **Future work: known-groups validity and Meaningful Within-Patient Change Estimation**