Missingness in Actigraphy Data – a case study

https://www.mdpi.com/2075-4426/12/5/749

Shyla Jagannatha
Statistics and Decision Sciences
Statistical Modeling and Methodology

Date: November 6th , 2023
Basel Biometric Society Annual Meeting

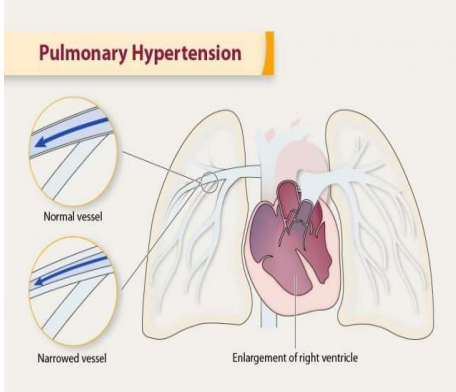
Johnson&Johnson
Innovative Medicine

Contents

- Brief introduction to design of the case study (TRACE)
- Visualizing actigraphy data, explore missingness
- Missing data mechanisms in studies with Actigraphy
- Feedback from FDA on the aggregation methods used in the defining the actigraphy endpoints
- Address FDA feedback and evaluate if the aggregation methods used in the TRACE study were acceptable

Why Actigraphy in Pulmonary Arterial Hypertension?

Pulmonary Arterial Hypertension (PAH) is a rare and debilitating disease affecting the pulmonary vasculature. Characterized by nonspecific symptoms such as reduced exercise capacity and dyspnea.



6-min walk test used to assess the exercise ability →

#6-Minute Walking test (6MWT) alone or as part of a composite

It is inexpensive, non-invasive, well-tolerated by patients and a valid measure of submaximal exercise capacity!

But, it is a snapshot and doesn't reflect the patient's daily life physical activity.

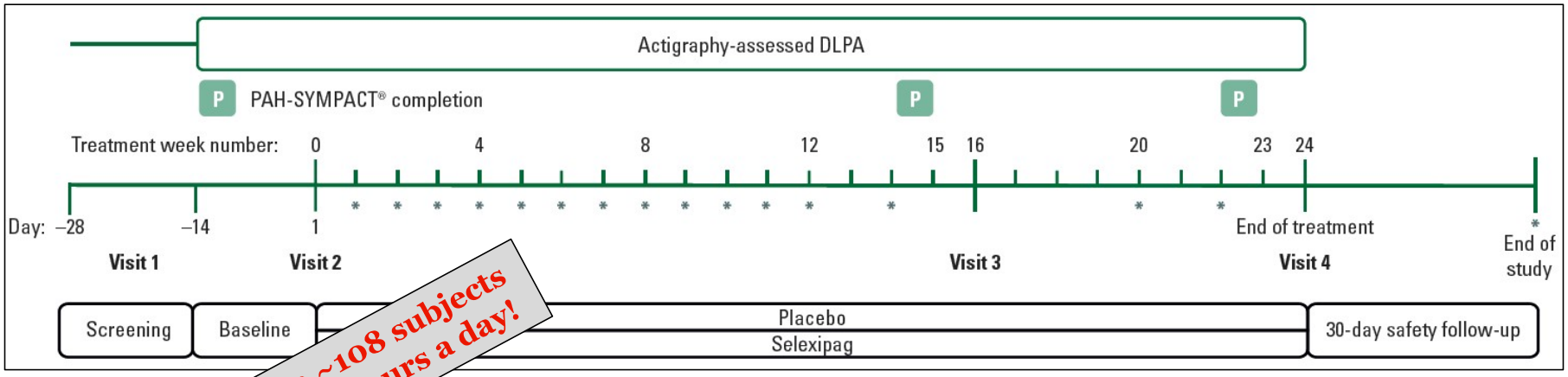
***Actigraphy:* Data on physical function in daily life may provide a more granular, patient-centric and accurate health related QoL (quality of life) measure to assess changes in clinical status.**

TRACE: Can actigraphy be used to assess changes in Daily Life Physical Activity (DLPA) in patients with PAH receiving selexipag or placebo?

*Continuing Cardiology Education, Volume: 4, Issue: 1, Pages: 2-12, First published: 27 July 2018, DOI: (10.1002/cce2.71)

Case Study: Study design of TRACE for Pulmonary Arterial Hypertension

- Prospective, multi-center, double-blind, placebo-controlled, Phase 4 study in patients with PAH to assess the effect of selexipag on Daily Life Physical Activity (DLPA) and QoL



Actigraphy data on ~108 subjects for ~168 days for 24 hours a day!

Howard LS, Rosenkranz S, Frantz RP, Hemnes AR, Pfister T, Hsu Schmitz SF, Skåra H, Humbert M, Preston IR. Assessing Daily Life Physical Activity by Actigraphy in Pulmonary Arterial Hypertension: Insights From the Randomized Controlled Study With Selexipag (TRACE). Chest. 2023 Feb;163(2):407-418. doi: 10.1016/j.chest.2022.08.2231. Epub 2022 Sep 8. PMID: 36089068; PMCID: PMC9899640.

Actigraphy data in TRACE

Physical activity using minute level (epoch) counts to define endpoints

→

All endpoints were to assess daily life activity:
Total number of Steps per day
Total of minutes in MVPA, etc.

→

Defined missingness in days.
For this, we need to define "a complete day" with sufficient wear time.
TRACE used **7 hours of "awake and wear"** as a **complete day/valid day**.

↓

Number of optimum "valid" days for reliable measure

TRACE used at least **7 "valid"** days for a **14-day period.**

←

Analyses by using daily aggregate of 14-day data assuming MAR

←

Compliance was quite high in the TRACE study.

Currently, "valid" day definition can be quite arbitrary.

There is some literature in specific populations or indications, but, wear hours and wear days vary widely.

Commonly used:
Wear time from 3 to 10 hours &
Wear days from 4 to 7 days

MVPA: Moderate to Vigorous Physical Activity
MAR: Missing at Random

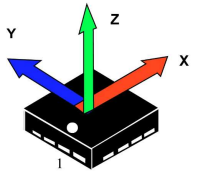
Actigraphy data

Minutes level Actigraphy data - one record per minute, per subject (>20 millions records)

A few records of one subject

Timestamp	Steps	AxisXCounts	AxisYCounts	AxisZCounts	SleepAwake	Wear
2018-11-16 09:51:00	0	117	168	343	AWAKE	TRUE
2018-11-16 09:52:00	0	361	176	189	AWAKE	TRUE
2018-11-16 09:53:00	0	0	0	0	AWAKE	TRUE
2018-11-16 09:54:00	0	0	0	0	AWAKE	TRUE
2018-11-16 09:55:00	0	0	0	0	AWAKE	TRUE
2018-11-16 09:56:00	0	0	0	0	AWAKE	TRUE
2018-11-16 09:57:00	0	0	0	0	AWAKE	TRUE
2018-11-16 09:58:00	0	0	0	0	AWAKE	TRUE
2018-11-16 09:59:00	0	0	0	0	AWAKE	TRUE
2018-11-16 10:00:00	0	0	0	0	AWAKE	TRUE
2018-11-16 10:01:00	0	0	0	0	AWAKE	TRUE
2018-11-16 10:02:00	0	0	0	0	AWAKE	TRUE
2018-11-16 10:03:00	0	0	0	0	AWAKE	TRUE
2018-11-16 10:04:00	0	660	931	1164	AWAKE	TRUE
2018-11-16 10:05:00	65	4651	3672	2468	AWAKE	TRUE

- **AxisXCounts** / **AxisYCounts** / **AxisZCounts**: Move intensity on the 3 orthogonal axes



(TOP VIEW)

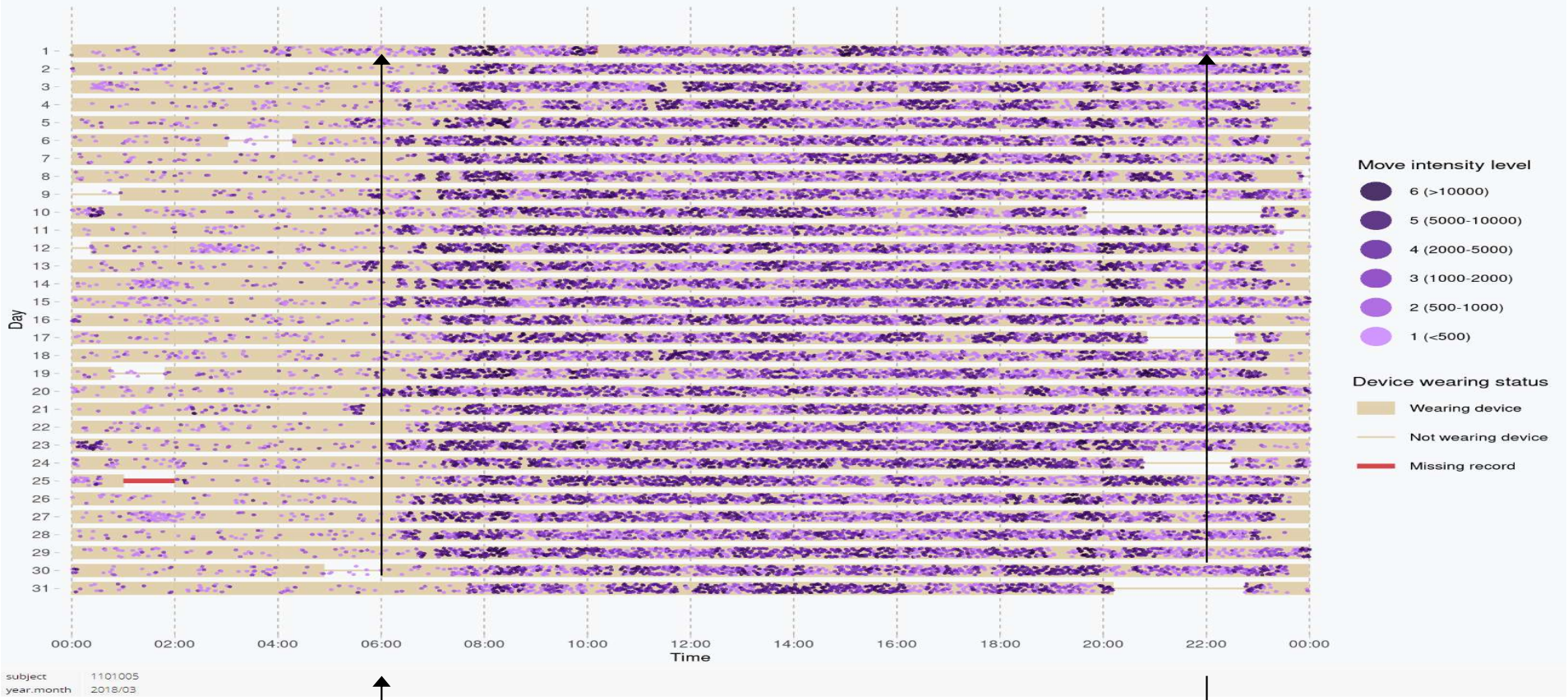
- **Steps**: Number of steps per minute
- **SleepAwake***: subject is considered as AWAKE or SLEEPING
- **Wear***: TRUE if subject is wearing the device, FALSE otherwise

* *Steps, SleepAwake, Wear are derived with Tudor-Locke/Troiano algorithms*

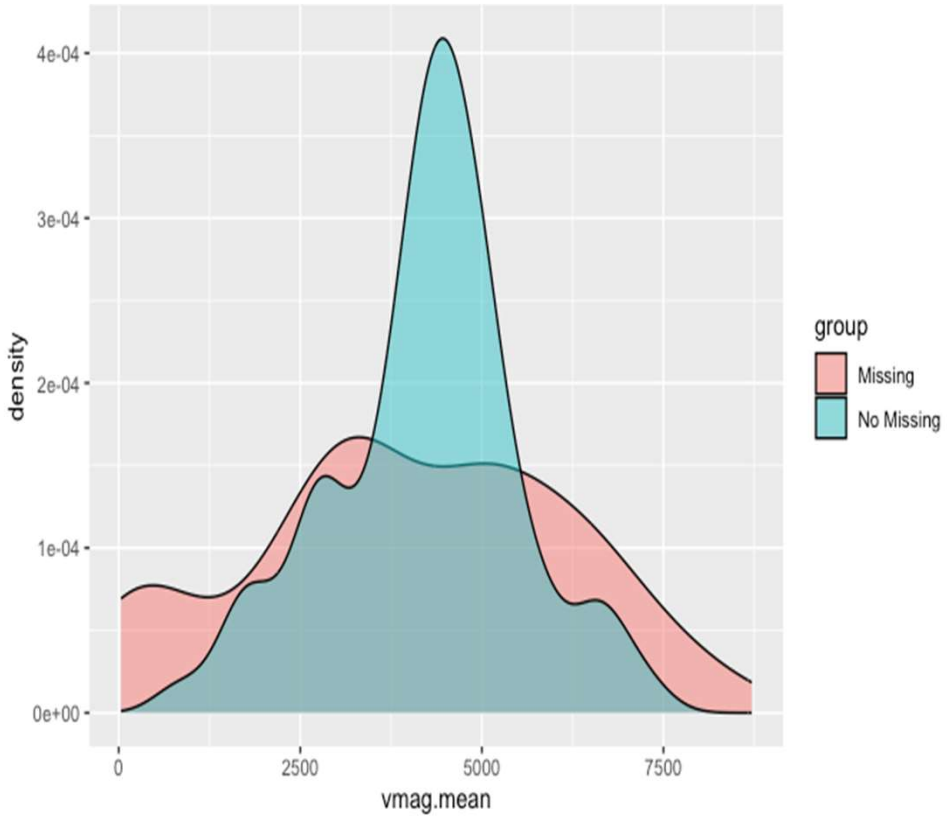
Vector magnitude $= \sqrt{(AxisXCounts^2 + AxisYCounts^2 + AxisZCounts^2)}$

Visualizing Actigraphy data (for one patient for 24 hours for 1 month)

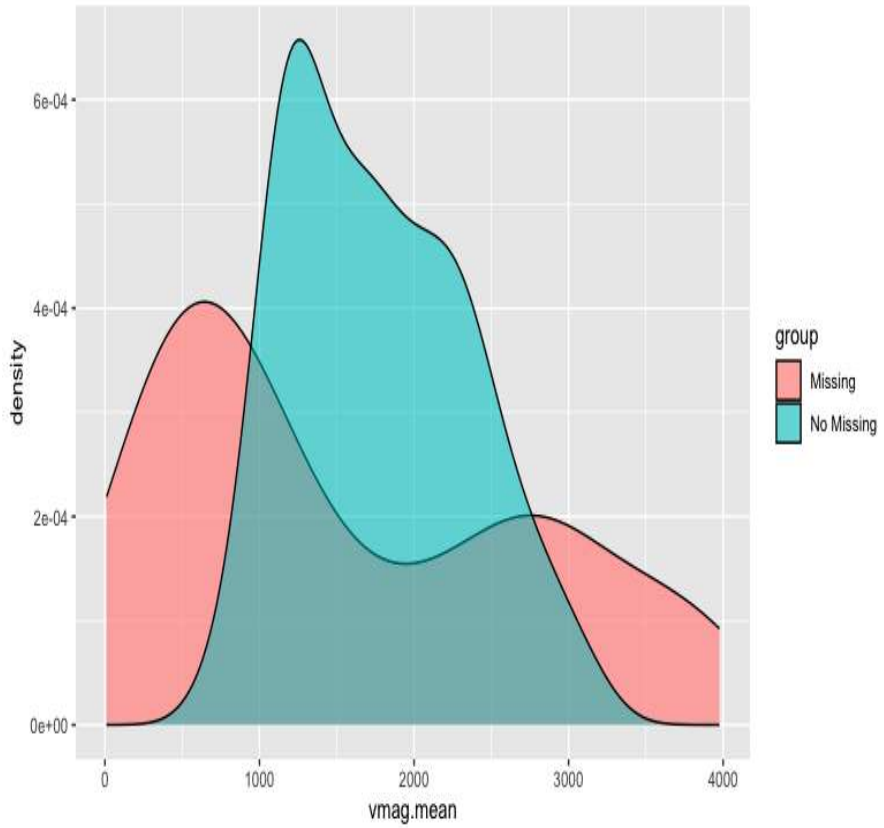
Wearing status with minute-level move intensity, One "row" is displaying a full day over 24 hours



Distribution of Daily Median Vector Magnitude of two patients for days that have no missing data and days that have missing data

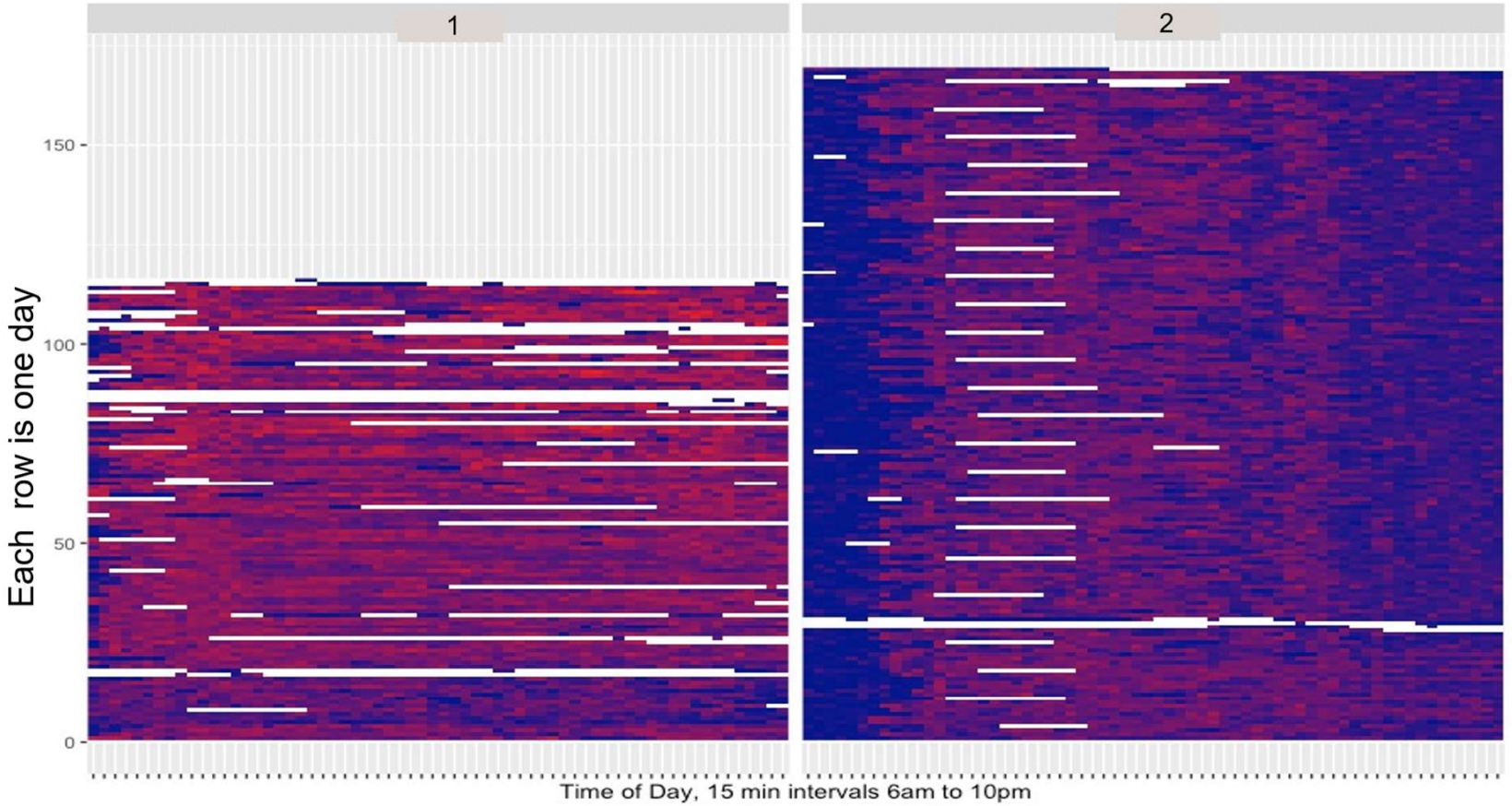


Daily Median for patient #1



Daily Median for patient # 2

Heatmap of physical activity over all days for two patients (Red is more active, blue of less active)



Median Vector Magnitude over days for patient #1

Median Vector Magnitude over days for patient # 2

A note on the missing data mechanisms in Accelerometry

Examples of Missing data mechanisms in Accelerometry data:

Missing Completely at Random (MCAR) Examples: Missing data due to device malfunction or data transfer error or due to charging, etc.

Missing at Random (MAR) Examples: Missing data more in a particular season, region, sex, disease severity, etc.

Missing Not at Random (MNAR) Examples: Patient fails to wear the device during severe symptoms, hospitalization, certain activities, etc.

More than one missing data mechanism can happen for a patient over the course of the trial.

The intercurrent events could be:

1. Misplaced Device
2. Change in attitudes, environment, season, etc.
3. If BYOD (Bring Your Own Device) is used, participants may change their device (upgrade version or a different brand)

All this leads to the question – what auxiliary data do we need to collect when using digital technologies in clinical trials?

Feedback from FDA on TRACE

- Justify if 7 hours is a representative sample of all activity during the day and if 7 days is representative sample of all daily activity during the 14-day period.
- Propose sensitivity analyses to explore the impact of this assumption on statistical inference
- Conduct additional analyses to explore missing completely at random (MCAR)

Steps to identify a "valid" day from TRACE study data

*Use days that had **no missing data** from **6 am to 10 pm (16 hours)***

Metric used for analyses: Vector Magnitude (VM):

Vector sum of counts measured in x, y, z axes:

VM (counts per minute) = $|(x, y, z)| = \sqrt{(x^2 + y^2 + z^2)}$

Induce missingness, **assuming MAR** based on each **patient's missing data patterns** (from days with missing data)

Calculate **median VM per minute** from a full day (no missingness), $\overline{VM}_{full}$ and for corresponding days with induced missingness, $\overline{VM}_{missing}$

Regress $\overline{VM}_{full}$ on $\overline{VM}_{missing}$ - separate model for each missing hours category:

$\overline{VM}_{full_{ij}} = \beta_0 + \beta_1 x_{ij} + u_j + e_{ij}$

Subject j, day i, x_{ij} is $\overline{VM}_{missing_{ij}}$ for missing hours 1, 2, 3, 4, …, 14

Calculate **percent bias (%*Bias*) in median VM per minute** in missing days relative to the full day

$\%Bias = \frac{(\overline{VM}_{full} - \overline{VM}_{missing})}{\overline{VM}_{full}} * 100$

Regress $\%Bias$ on missing hours category (1 to 14) :

one model regressing %Bias on duration of missing hours:

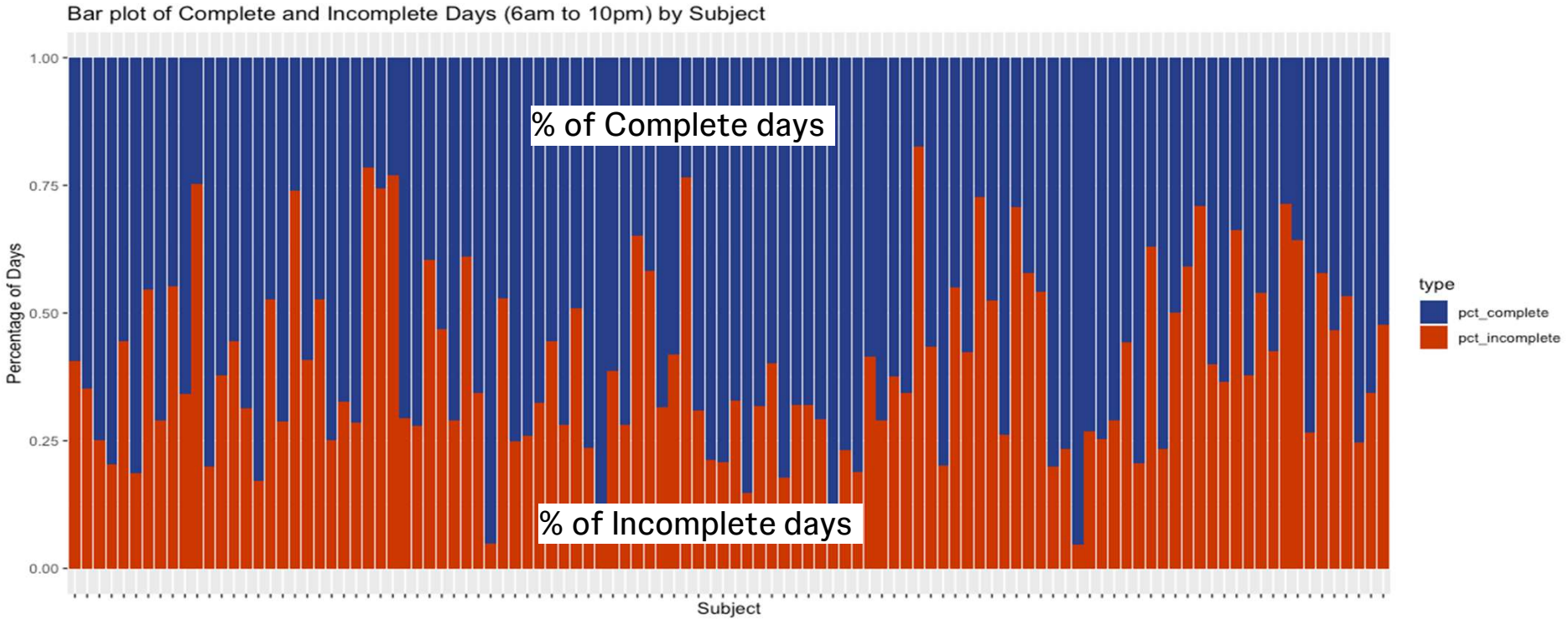
$\%Bias_{ijk} = \beta_0 + \beta_1 x_{ijk} + u_j + e_{ijk}$

Subject j, day i, x_{ijk} is k^{th} categorical missing hours (1, 2, …, 14) for subject j, day i,

Step1: Induce missingness (Patient's own missing pattern method)

Each patient has x% of complete days and 1-x% of incomplete days

- Obtain the missing data patterns for each patient using their incomplete days
- Induce missingness, based on each **patient's missing data patterns** (from days with missing data) using *ampute function in {mice} package*

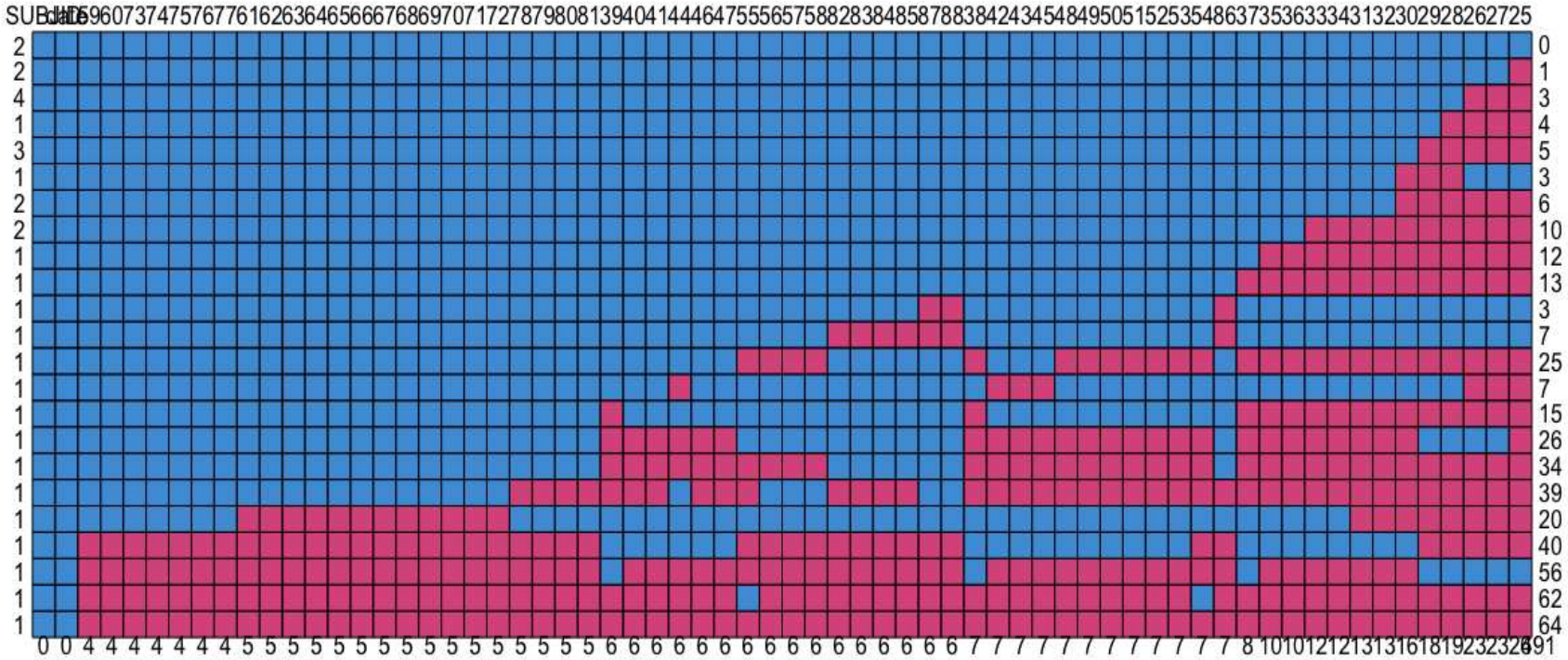


Missing data pattern for one patient x from incomplete days

Each square/pixel is one 15 min window between 6am to 10pm for a day with missing data

Column: column1 = Subject ID, column2 = Day, Columns 3-64 = 15 min windows from 6am to 10pm

Row: Each row is a pattern of missingness

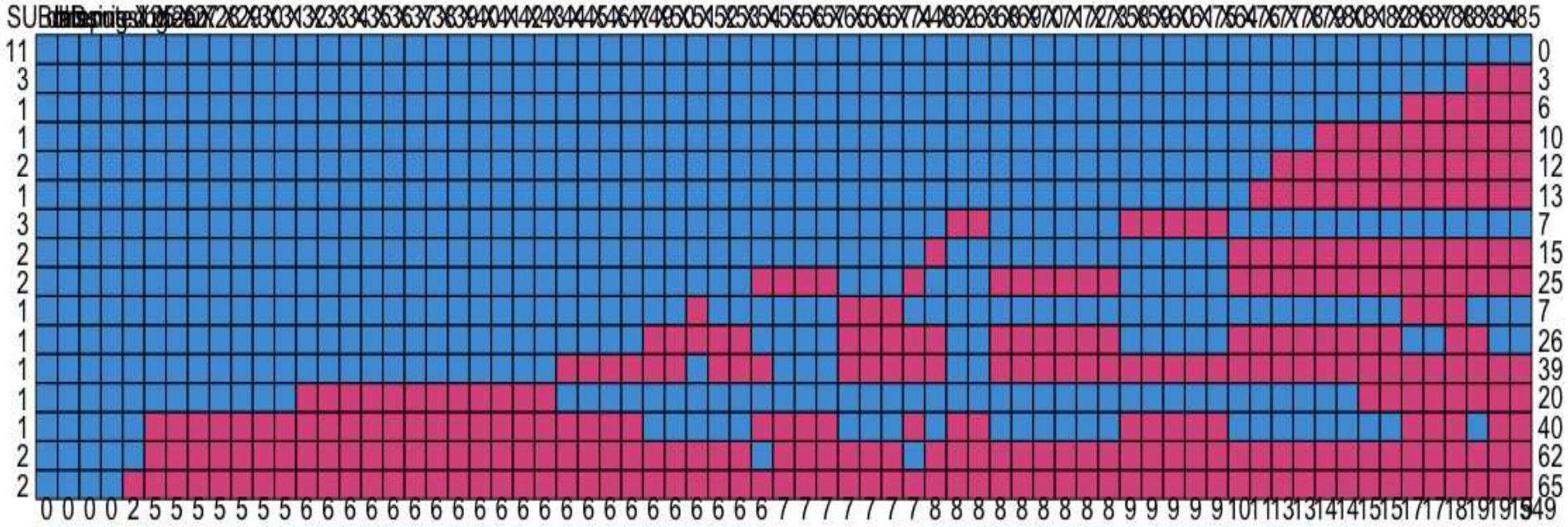


Amputed missing data pattern for one patient x from complete days

Each square/pixel is a one 15 min window between 6am to 10pm for a day with missing data

Column: column1 = Subject ID, column2 = Day, Columns 3-64 = 15 min windows from 6am to 10pm

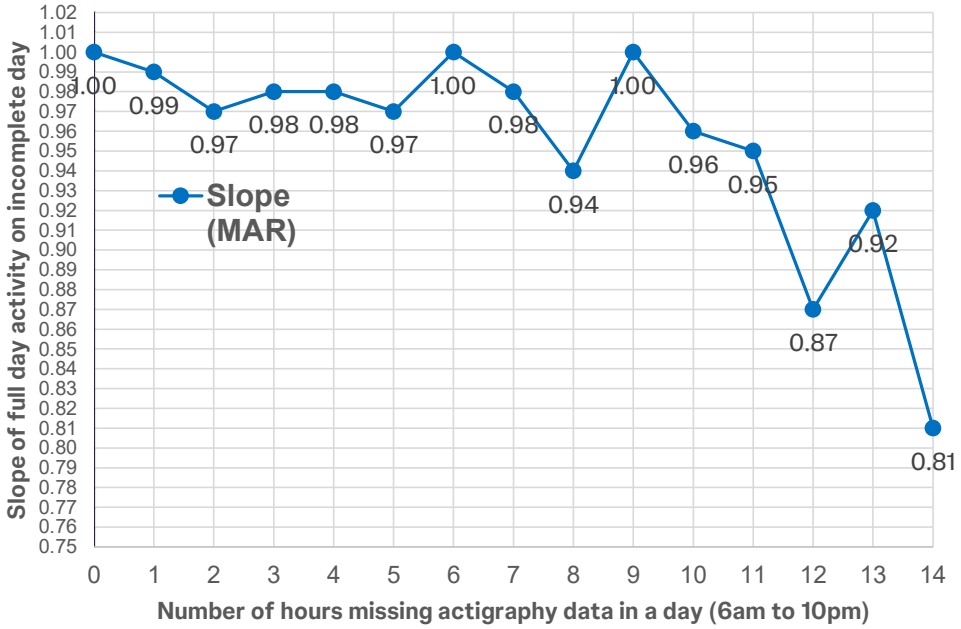
Row: Each row is a pattern of missingness



Step 2: Evaluate the induced missingness (missing pattern method)

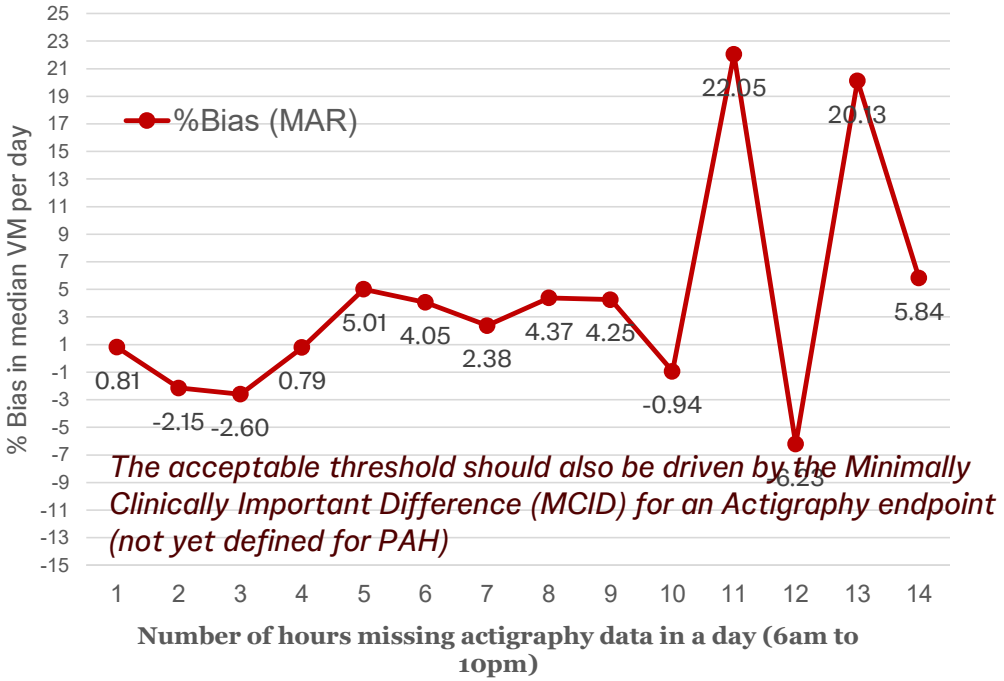
Results of Regression of $\overline{VM}_{full}$ and % Bias on $\overline{VM}_{missing}$

Slope of $(VM)\_full$ day on (VM)_incomplete day



For a slope to be <0.95, missing hours should be ≤ ~8 hours, i.e. ~ 8 hours of wear time.

% Bias for incomplete days relative to full day



The acceptable threshold should also be driven by the Minimally Clinically Important Difference (MCID) for an Actigraphy endpoint (not yet defined for PAH)

%Bias : bias of the missing day estimate compared to a full day. *Lower the better.*
For a bias of ≤ 5%, missing hours should be ≤ 10 hours, i.e. ~6 hours of wear time.

Did we justify if 7 hours of awake and wear time is sufficient to capture a PAH patient's daily physical activity?

- We used one simple metric. This needs to evaluated for other endpoints.
- It may be more realistic to use the patient's own missing data patterns if feasible.
- The acceptable thresholds for bias may depend on the minimally clinically meaningful change for that endpoint.
- The missingness was MAR. The thresholds may be different if other types of missingness are mixed in.

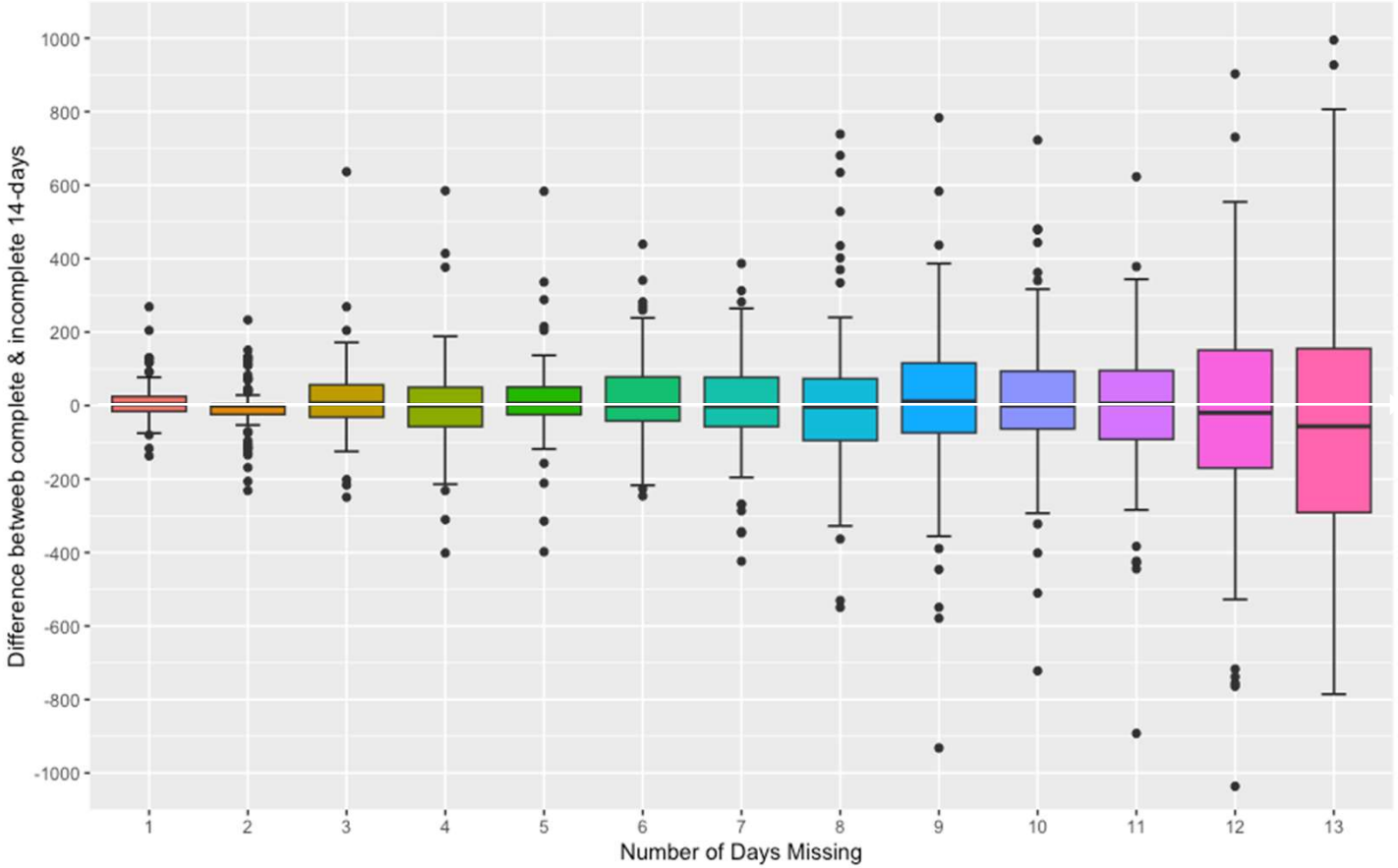
- *Given the above caveats, for this one simple metric, 6-8 hours of awake and wear time gives a reasonable estimate of the daily activity.*

Justify if 7 days is a representative sample of a 14-day window

- Identify patients who had a minimum 7 hours of wear data during the day for 2 weeks
- 97/108 patients met this criteria for the first 2 weeks
- Induce missing days (from 1 to 13 days) for those patients with complete 14 days
- Compare the bi-weekly medians between “complete 14 days” and “incomplete 14 days”

Justify if 7 days is a representative sample of a 14-day window

Box plot of difference between medians from complete 14-days and incomplete 14-days (along the X-axis indicating the number of missing days)



1. As the number of missing days increase, the variation in the estimated median increases as well (expected).
2. Missing 12 or more days introduces bias in the estimate.

So, how many days of data are a representative sample of a 14-day window?

Minimum number of wear days to obtain reliable data Spearman-Brown Prophecy formula

Intraclass correlation coefficient (of daily data for a subject)

ICC using amputed data with 7 hours of wear time = 0.44 & using the average of amputed data (from 4 to 15 hours wear time) = 0.55

To calculate the minimum number of wear days to obtain a reliable estimate, Spearman-Brown Formula was used.

The Spearman-Brown Formula (or the Spearman-Brown Prophecy Formula) is a **measure of test reliability**. It's usually used when the **length of a test** is changed and its effect on reliability.

The formula is:

$r_{kk} = k(r_{11}) / [1 + (k - 1)* r_{11}]$

Where:

- r_{kk} = reliability of a valid day for "k" days,
- r_{11} = reliability of a single day
- k = factor by which the length of the valid days is changed.

	Reliability for wear days								
Single Day Reliability	2 days	3 days	4 days	5 days	6 days	7 days	8 days	9 days	10 days
0.44 (for 7 hours daytime wear)	0.61	0.70	0.76	**0.80**	0.83	0.85	0.86	0.88	**0.89**
0.55 (using average)	0.71	0.79	**0.83**	0.86	0.88	0.89	**0.90**	0.91	0.92

Concluding Notes:

- The thresholds used in the TRACE study were justifiable using a generic simple measure! The methods and thresholds have to be evaluated for sensitivity and re-confirmed for a different measure or population.

- Currently:
  - There is an explosion of digital health technology that have potential for precise and timely assessments of health endpoints.
  - We are in the stages of developing and validating the endpoints.
- Near Future:
  - These endpoints may become primary and secondary endpoints in registrational clinical trials.
  - The questions around what type of data to collect, how to define, address and impute missing data will become pertinent for statisticians.
  - We need to start having such discussions now.

Acknowledgements for support and collaboration

- Mathias Lebreton
- Matthieu Villeneuve
- Yoko Shiraga
- Marisa Bacchi
- Brian Hennessy
- Ciprian Crainiceanu

References:

- Howard LS, Rosenkranz S, Frantz RP, Hemnes AR, Pfister T, Hsu Schmitz SF, Skåra H, Humbert M, Preston IR. Assessing Daily Life Physical Activity by Actigraphy in Pulmonary Arterial Hypertension: Insights From the Randomized Controlled Study With Selexipag (TRACE). Chest. 2023 Feb;163(2):407-418. doi: 10.1016/j.chest.2022.08.2231. Epub 2022 Sep 8. PMID: 36089068; PMCID: PMC9899640.
- TROST, STEWART G.1; MCIVER, KERRY L.2; PATE, RUSSELL R.2. Conducting Accelerometer-Based Activity Assessments in Field-Based Research. Medicine & Science in Sports & Exercise 37(11):p S531-S543, November 2005. | DOI: 10.1249/01.mss.0000185657.86065.98
- Rianne Margaretha Schouten, Peter Lugtig & Gerko Vink (2018) Generating missing values for simulation purposes: a multivariate amputation procedure, Journal of Statistical Computation and Simulation, 88:15, 2909-2930, DOI: 10.1080/00949655.2018.1491577
- Schouten, R. M., & Vink, G. (2021). The Dance of the Mechanisms: How Observed Information Influences the Validity of Missingness Assumptions. Sociological Methods & Research, 50(3), 1243-1258. https://doi.org/10.1177/0049124118799376
- Actigraphy as a clinically meaningful endpoint to detect change after treatment with iNO (30 mcg/kg-IBW/hr) in patients at risk of Pulmonary Hypertension associated with Pulmonary Fibrosis . S.D. Nathan1, K. Flaherty2, M. K. Glassberg3, G. Raghu4, J. Swigris5, R. Alvarez3, N. Ettinger6, J. Loyd7, P. Fernandes8, H. Gillies8, P. Shah8, L. Lancaster7 , https://bellerophon.gcs-web.com/static-files/e27eea3d-de6c-4f30-abf8-f28ec0eb6550
- Demeyer H, Burtin C, Hornikx M, Camillo CA, Van Remoortel H, Langer D, et al. (2016) The Minimal Important Difference in Physical Activity in Patients with COPD. PLoS ONE 11(4): e0154587. https://doi.org/10.1371/journal.pone.0154587

Thank you

Time for questions/feedback