Early experiences using SDTM to organize clinical data collected using a physical activity tracker

PHUSE EU CONNECT 2019
Martin Gram and Gianluca Mortari
Introduction

- Real World Evidence provides improved insight
- Wearable technologies play a growing role in clinical trials
- We can pursue more readily useful endpoints
- One application is to investigate the physical activity level

- However, it is not clear how to incorporate device data in clinical trials
Select the device

• Consumer grade devices
  • Apple Watch, Fitbit, Garmin etc.
  • Proprietary algorithms
  • Data transparency low
  • Data validity unknown

• Research grade devices
  • Digiwalker, Actigraph, Omron etc.
  • Publicly available algorithms
  • Data transparency high
  • Data validity can be documented

Other things to consider:
• FDA classified (Class I, II or III)
• Sufficient battery life
• Simple design, robust, waterproof
• Continuous data upload to cloud
The Anatomical Wear Location

- Wrist
- Hip
- Ankle
- Thigh
- Lower back
- Biceps
The Anatomical Wear Location

Wrist placement
- ! Current algorithms less mature
- ✓ Reported non-wear lower
- ✓ Enhances wear compliance
- ✓ Wear comfort is high

Hip placement
- ✓ Current algorithms more mature
- ✓ Near center of gravity
- ! Report non-wear higher
- !! Patient burden higher

Which anatomical location is considered valid in a submission?
Non-Wear and Sleep

A normal day...

Example Counts

<table>
<thead>
<tr>
<th>Time</th>
<th>Wear</th>
<th>Non-wear</th>
<th>Wear</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>3000</td>
<td>3000</td>
</tr>
<tr>
<td>3</td>
<td>500</td>
<td>0</td>
<td>3000</td>
</tr>
<tr>
<td>6</td>
<td>500</td>
<td></td>
<td>3000</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>250</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sleep

Time
Non-Wear

- Non-wear are periods during a day where the subject is not wearing the device.
- In real life subjects will take the device off during a trial.
- Zero activity is recorded during this period although the subject likely did some activity.
- It follows that any non-wear during waking hours will underestimate the total daily activity level.
Approaches to handle non-wear

- Ignore it
- Extrapolate the activity pattern from the other part of the day
- Extrapolate the activity pattern from other similar days (time-specific imputation)
  - Multiply with the same number of hours for all subjects
  - Multiply with the number of waking hours
- Only accept days having very high wear time
- Report the data as percentage of wear time

How is non-wear handled in the best way?
Sleep

- If physical activity is the focus, sleep would usually be excluded
- Sleep can be excluded manually or automatically using:
  - Individualised time filters
  - Standard time filters

Challenge:
- Sleep, sedentary activity and non-wear looks similar
- Classification of awake sedentary time and movement during sleep is difficult

How is sleep handled in the analysis?
Availability of Fit-For-Purpose Algorithms

These algorithms have not been validated in the full age range. Is the data output considered valid in a submission?
When is an algorithm validated, and when is it good quality?

A step...
Easy and intuitive to understand, and often used...

...or is it:

**step**  [step]

*noun*

1. a movement made by lifting the foot and setting it down again in a new position, accompanied by a shifting of the weight of the body in the direction of the new position, as in walking, running, or dancing.
Some challenges with steps

- What is a step exactly?
- Difficult to measure on the wrist
- Prone to type 1 error
- Activity and gait pattern different between children and adults
- Very few validations done in free-living conditions
- An algorithm validated on all criteria (age, anatomical location, population and the endpoint steps) can still be inaccurate

Several “validated” step algorithms exist. Arguably none of them are accurate in a free-living environment
Data granularity

- The collection of minute level data is challenging on many aspects
  - Data volume
  - Processing time
  - Submission
  - ...

- But scaling up the data granularity, could potentially mask interesting patterns hidden in the data

Activity data collection

- All collected data must be tabulated in SDTM format
- About 300 subjects enrolled
- 12 Weeks (>120,000 Minutes)
- 25 Derived parameters

>800,000,000 records
>200 gb
SDTM domains selection

- **XA – DX**
- **XA – SUPPXA (UTC Datetime)**
- **DX – DI (SPDEVID)**

### USUBJID, XAGRPID, XATESTCD, XARRES, XARRESU, XADTC

<table>
<thead>
<tr>
<th>USUBJID</th>
<th>XAGRPID</th>
<th>XATESTCD</th>
<th>XARRES</th>
<th>XARRESU</th>
<th>XADTC</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNXXXX-YYYYY/123456</td>
<td>86</td>
<td>AC11WEAR</td>
<td>Y</td>
<td></td>
<td>2019-07-25T14:12</td>
</tr>
<tr>
<td>NNXXXX-YYYYY/123456</td>
<td>86</td>
<td>AT14SLA</td>
<td>N</td>
<td></td>
<td>2019-07-25T14:12</td>
</tr>
<tr>
<td>NNXXXX-YYYYY/123456</td>
<td>86</td>
<td>AT18SLAD</td>
<td>N</td>
<td></td>
<td>2019-07-25T14:12</td>
</tr>
<tr>
<td>NNXXXX-YYYYY/123456</td>
<td>86</td>
<td>S15ADLI</td>
<td>30 sec</td>
<td>sec</td>
<td>2019-07-25T14:12</td>
</tr>
<tr>
<td>NNXXXX-YYYYY/123456</td>
<td>86</td>
<td>S15ADMO</td>
<td>30 sec</td>
<td>sec</td>
<td>2019-07-25T14:12</td>
</tr>
</tbody>
</table>

### USUBJID, SPEVID, DXLAT, DXSTDT, DXENDT

<table>
<thead>
<tr>
<th>USUBJID</th>
<th>SPEVID</th>
<th>DXLAT</th>
<th>DXSTDT</th>
<th>DXENDT</th>
</tr>
</thead>
</table>

### ActivityData/DDDDMMYYYY.zip

- **DI.csv**
- **DX.csv**

### Supplimentary Data

- **USUBJID, IDVARVAL, QNAM, QVAL**
  - NNXXXX-YYYYY/123456: 86 TIMEZONE America/New_York

- **SPDEVID, DIPARMCD, DIVAL**
  - XYZ1A37180280 TYPE Kinesiology Ambulatory Recorder
  - XYZ1A37180280 MANUF Provider Name
  - XYZ1A37180280 MODEL XYZ01
  - XYZ1A37180280 SERIAL XYZ1A37180280
  - XYZ1A37180280 FDAUDI FDAUDI Number
## Parameters categorisation

<table>
<thead>
<tr>
<th>Number of TESTCD’s</th>
<th>ALGORITHM</th>
<th>ENDPOINT</th>
<th>ADOLESCENTS</th>
<th>ADULTS</th>
<th>SEDENTARY</th>
<th>LIGHT</th>
<th>MODERATE</th>
<th>VIGOROUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CHOI 2011</td>
<td>WEAR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>AXIS COUNTS</td>
<td>ACTIVITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>TRACY 2014/2018</td>
<td>SLEEP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>UWF STEPS</td>
<td>ACTIVITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>STAUDENMAYER 2015</td>
<td>ACTIVITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>VA CROUTER 2015</td>
<td>ACTIVITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>VM CROUTER 2015</td>
<td>ACTIVITY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Relevance of SDTM variables

- **Preset variables**
- **Duplicated variables**
- **Probably unnecessary variables**

<table>
<thead>
<tr>
<th>STUDYID</th>
<th>USUBJID</th>
<th>XGRPID</th>
<th>XATEST</th>
<th>XAREASND</th>
<th>XASTAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOMAIN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>XAGRPID</td>
<td></td>
<td>XACAT</td>
<td>XABLFL</td>
<td>XEREASND</td>
</tr>
<tr>
<td></td>
<td></td>
<td>XACAT</td>
<td>XASCAT</td>
<td>XADTC</td>
<td>XASTAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>XASTESC</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>XORRES</td>
<td>XASTRESN</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>XORRESU</td>
<td>XASTRESU</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>XETCD</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EPOCM</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>XADY</td>
<td></td>
</tr>
</tbody>
</table>
Questions

- Which anatomical location is considered valid in a submission?
- How is non-wear and sleep handled in the best way?
- When is an algorithm validated, and when is it good quality?
- Are all expected variables in SDTM necessary in activity data?
- Will alternative naming conventions for split domains be considered valid in a submission?
- Is SDTM following the evolution of the data?
Conclusions

- Wearable medical devices have a great potential
- Work still needs to be done to solve challenges on data collection
- Uncertainty on submission
- Building awareness, kickstarting discussion

xmgm@novonordisk.com
glcm@novonordisk.com
Thank you for the attention