SciKit Digital Health: Python Package for Streamlined Wearable Inertial Sensor Data Processing

October 5, 2022 | 11am ET
But first, housekeeping

• Please note today’s session is being recorded
• To ask a question for discussion during Q&A, please:
  • Either ‘raise your hand’ in the participant window and moderator will unmute you to ask your question live, or
  • Type your question into the chat box
• Slides and recording will be available after today’s session
Virtual Journal club

SciKit Digital Health: Python Package for Streamlined Wearable Inertial Sensor Data Processing

October 5, 2022 | 11am ET
# SciKit Digital Health

## What does SKDH address?

### Existing Solutions & Literature
- **Lots** of literature on gait, activity, sleep, etc
- **Lack** of existing **implementations** from literature
- **Lack of Open-Source** algorithms or implementations
- Packages do exist, but have drawbacks
  - GGIR: sleep & activity only, not easily extensible
  - GaitPy (Pfizer): gait only, harder to interface
- Many packages are one area only (e.g. activity only, or gait only, etc.)

### Python Algorithms Package
- **Distributable** package in common language
- Open-Source
- Algorithms from literature & internal research
- Algorithms **validated** against ground truth
- Consolidate common algorithms into one package/repository
- Custom framework handles multiple pipeline variations
- Clean & well documented code
## Scikit Digital Health – General Purpose Modules

### IO

- Read common (binary) file types into memory for processing
  - GeneActiv
  - Axivity
  - APDM
  - Etc.
- Low-level language extensions for speed

### Signal Features

- Suite of common signal features
- Common framework allows easy extension/custom features
  - Can be computed at same time as rest of built-in features
- Low-level language extensions for speed

### Pre-Processing

- Pre-processing algorithms for multi-day accelerometer data
- Wear detection (van Hees 2013)
- Accelerometer calibration (van Hees 2014)

---


**Scikit Digital Health – Gait Algorithm & Model**

### Algorithm Details

- **Lumbar accelerometer based**
- **Provides temporal (e.g. stride time) & spatial (e.g. stride length) gait metrics**

**Algorithm details:**
- Custom classifier detects gait bouts (3s windows)
- Continuous wavelet transform (CWT) based initial/final contact detection (McCamley 2012)
- Inverted pendulum model for spatial metrics (Zijlstra 2003)

**Quality check per step/strides:**
1. Loading time less than P% of max. stride time (20%)
2. Stance time less than half gait cycle + initial double support
3. Maximum stride time not exceeded (2.25s)

- Computers signal-based asymmetry endpoints

---


Scikit Digital Health – Gait Validation & Applications

Gait Validation

Example Applications

• Healthy Adults:
  • Showed good-excellent agreement with GaitRite mat
  • Significant differences between young (18-40) and old (65-80) groups at-home
  • Minimum monitoring durations of ~3-4 days

• Healthy Children:
  • Adjust expected step frequency (CWT IC/FC detection) and maximum stride time
  • Good-excellent agreement with GaitRite mat

• Children with Achondroplasia

• Adults with knee osteo-arthritis


### Algorithm

- **Wrist accelerometer based**
- **Computes MVPA, sedentary/light time, max accel. over time windows**
- **Algorithm details:**
  - Compute Euclidean Norm Minus One (ENMO) and window (default 5s)
  - Thresholds from previous literature for estimating activity level duration
  - Additionally computes other endpoints:
    - Intensity Gradient (Rowlands 2018)
    - Fragmentation endpoints (Di 2017)
  - Framework for adding other custom endpoints

### Example Applications

- **Healthy Adults:**
  - [In-progress] Agreement with other open/closed source algorithms
    - ActiGraph, GeneActiv, & GGIR
  - [In-progress] Investigate group differences between younger (18-40) and older (65-80) participants at-home
- **Cachexia**
- **Heart Failure**
- **etc**

---

Scikit Digital Health – Sit-to-Stand

Algorithm

- Lumbar based accelerometer-only
- Computes duration, STS dynamics, signal smoothness
- Algorithm details:
  - Orientation independent
  - Heuristic algorithm uses CWT to detect potential sit-to-stand transfers
  - Each possible location is QC’ed to eliminate partial or false transfers

Validation & Application

- Validated in healthy and patients with Parkinson’s Disease (Adamowicz 2020)
- Ongoing work in people with knee OA, cachexia

Scikit Digital Health – Sleep

Algorithm

• Wrist accelerometer based
• Computes sleep duration, sleep fragmentation, awake time
• Algorithm details:
  • Updated implementation of SleepPy (Christakis 2019)
  • Heuristic algorithms based on activity measures and estimated arm-angle

Application & Validation

• Validation against PSG in people with Atopic Dermatitis (Mahadevan 2019)
• [In-progress] Healthy populations

Scikit Digital Health – Package Framework

- **Common input/output** → chain process modules
- **Pipeline structure** → simple expression of complex/arbitrarily-ordered processing steps
  - Pipeline: processing steps for generating DHEs
- User-adjustable algorithm parameters
- Pipelines can be saved & loaded for repeatability or documentation
- **Common structure** → extensible
Creating a pipeline instance

```python
import skdh

pipeline = skdh.Pipeline()
pipeline.add(skdh.io.ReadBin())
pipeline.add(skdh.preprocessing.CalibrateAccelerometer())
pipeline.add(skdh.preprocessing.DetectWear())
pipeline.add(skdh.sit2stand.Sit2stand())
pipeline.add(skdh.gait.Gait())

pipeline.run(file="example_geneactiv_file.bin")
```
Scikit Digital Health – Extending

**Ease of Extending**

- Public base class
- Create custom/one-off extensions for niche cases
- Mix & match SKDH and extensions in pipelines
Scikit Digital Health – Documentation


Background Information

General terminology:
- Initial Contact (IC): the first contact of a foot with the ground, also "Heel Strike"
- Final Contact (FC): the last contact of a foot with the ground, also "Toe Off"
- Stride: between ICs of the same foot, eg IC(i) to IC(i+2), IC(i+1) to IC(i+3)

Per stride terminology:
- Step: between ICs of opposite feet, eg IC(i) to IC(i+2), IC(i+1) to IC(i+2). There are 2 steps to every stride
- Stance: when the foot is in contact with the ground, eg IC(i) to FC(i)
- Swing: when the foot is not in contact with the ground, eg FC(i) to IC(i+2)
- Double Support: when both feet are in contact with the ground simultaneously
- Single Support: when only 1 foot is in contact with the ground

Adding Custom Gait Endpoints

A modular system for computing gait endpoints is employed to aid in the addition of custom gait endpoints. Two base classes exist depending on what type of endpoint is being added:
Scikit Digital Health - Accessing

https://github.com/PfizerRD/scikit-digital-health

https://anaconda.org/conda-forge/scikit-digital-health

https://pypi.org/project/scikit-digital-health/
THANK YOU

Yiorgos and Lukas!

Acknowledgements:

- Isik Karahanoglu
- Junrui Di
- Wenyi Lin
- Charmaine Demanuele
- Pfizer DSTI Colleagues
- Tomasz Adamusiak* (former Pfizer Employee)
- Matt Czech* (former Pfizer Employee)
Should Alexa diagnose Alzheimer’s?: Legal and ethical issues with at-home consumer devices

November 2nd 11am EST
### Scikit Digital Health: Gait Endpoints

<table>
<thead>
<tr>
<th>Metric</th>
<th>GaitPy [2]</th>
<th>SKDH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stride/Step Time/Cadence</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Stride Phases</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Single/Double Support</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Stride/Step length</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Gait Speed</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Basic Asymmetry</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Intra-stride/step-covariance</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Harmonic Ratio</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Stride SPARC</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Regularity Indices/GSI</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Phase Coordination Index</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Autocovariance Symmetry</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>


https://doi.org/10.21105/joss.01778
### Scikit Digital Health: Activity Endpoints

<table>
<thead>
<tr>
<th>Metric</th>
<th>SKDH</th>
<th>GGIR</th>
<th>GENEActiv</th>
<th>Actigraph</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sed/Light/Mod/Vig Duration</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Bouted activity level Duration</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MET estimates</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Intensity Gradient (Rowlands 2018)</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max. Accel. in windows</td>
<td>x</td>
<td></td>
<td>Computable</td>
<td>Computable</td>
</tr>
<tr>
<td>Avg. Duration</td>
<td>x</td>
<td></td>
<td>Future?</td>
<td></td>
</tr>
<tr>
<td>State Transition Prob.</td>
<td>x</td>
<td></td>
<td>Future?</td>
<td></td>
</tr>
<tr>
<td>Gini Index</td>
<td>x</td>
<td></td>
<td>Future?</td>
<td></td>
</tr>
<tr>
<td>Avg. Hazard</td>
<td>x</td>
<td></td>
<td>Future?</td>
<td></td>
</tr>
<tr>
<td>State Power Lab Dist.</td>
<td>x</td>
<td></td>
<td>Future?</td>
<td></td>
</tr>
</tbody>
</table>
### Scikit Digital Health: Sit-to-Stand Endpoints

<table>
<thead>
<tr>
<th>Metric</th>
<th>SKDH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration</td>
<td>x</td>
</tr>
<tr>
<td>Vertical Displacement</td>
<td>x</td>
</tr>
<tr>
<td>Max. Acceleration</td>
<td>x</td>
</tr>
<tr>
<td>Min. Acceleration</td>
<td>x</td>
</tr>
<tr>
<td>Spectral Arc Length</td>
<td>x</td>
</tr>
</tbody>
</table>
### Scikit Digital Health: Sleep Endpoints

<table>
<thead>
<tr>
<th>Metric</th>
<th>SKDH</th>
<th>SleepPy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sleep Time</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Percent Time Asleep (sleep efficiency)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Number of Wake Bouts</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Sleep Onset Latency</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Wake After Sleep Onset</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Average Sleep Duration</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Average Wake Duration</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Sleep-Wake Transition Probability</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Wake-Sleep Transition Probability</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Sleep Gini Index</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Wake Gini Index</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Sleep Average Hazard</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Wake Average Hazard</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Sleep Power Law Distribution</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>Wake Power Law Distribution</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>
Gait in Healthy Adults (23-39 & 65-85 years)

Scikit Digital Health – Sit-to-Stand At-Home

Minimum monitoring duration for Sit-to-Stand transfer endpoints

Key Points

- After 2 days all median endpoints have median ICC > 0.7
- After 5 days all endpoints have median ICC > 0.9