Data-driven Digital Healthcare:
Developing Effective Digital Biomarkers for Parkinson’s Disease

George Roussos
g.roussos@bbk.ac.uk
Overview

- cloudUPDRS certified Class I Medical Device for clinical use
- Rate PD motor symptoms as precisely as an experienced clinician
- Extends and adapts Part III of the standard UPDRS protocol
- Unsupervised use at home
  - Employs accelerometer for tremor and gait measurements
  - Employs touch-screen for tapping measurements
  - No clinical or technical supervision during testing: bespoke user journey
- Data analytics
Challenges

1. Ensure unsupervised test is carried out correctly
2. Reduce testing time
3. Capture symptom variability
4. Identify high quality signal segments
People with Parkinson’s
Parkinson’s Disease (PD)

- No cure
- Managed mainly by replacing dopamine
- Motor symptoms
  - tremor, rigidity, slowness of movement (bradykinesia), freezing of gait, stiffness, shaking, falls
- Non-motor symptoms
  - bladder, memory, sleep, addictive behaviour, fatigue, pain, hallucinations
Symptoms of Parkinson’s Disease

resting tremor
parkinson’s disease
drcrunch.co.uk
Disease Progression

- Prodromal: Sleep disorder, Loss of smell, Constipation, Depression
- Early: Slowness, Stiffness, Shaking, Urgency
- Moderate: Anxiety, Low blood pressure, Dyskinesias, Impulsivity, Freezing
- Complex: Brittle ON/OFF’s, Cognitive impairment, Pain, Hallucinations

Impairment over time (years)

Levodopa response

- Levodopa
- MAOI
- Agonists

- COMT Agonists

- Deep brain stimulation
- Apomorphine
- Duodopa

Treatments
- Standard clinical protocol for assessing PD
- Part III clinical assessment of motor symptoms
- Known issues:
  - time intensive
  - inter-rater variability
  - not sensitive
- Used formally
  - Drug trials
  - Consideration of advanced therapies
- Used informally as part of clinical assessment (once or twice per year) of disease progression
- Can we replace with an app?
Expressed use intentions

- Quantitative and qualitative methods
  - survey and audience panels
- Majority of PD patients would use app (86%)
  - Most would prefer the test to last less than 5 minutes per assessment (64%)  
  - Some would accept up to 10 minutes (27%)
- Main motivation: Need to understand their condition
- No expressed privacy concerns
Information Processing Pipeline

Smartphone apps
- Modality: batch processing
- Format: JSON, plain TXT
- Data converted to standardized representation based on PANDAS dataframes
- Ingestion is extensible so any data format can be loaded

• Feature extraction
  - 500+ tremor, bradikynesia, tapping, gait, turning and speech features
  - Implemented in the TremorProcessor, FingerTappingProcessor, GaitProcessor, SpeechProcessor classes

• Supervised and unsupervised learning for scoring
  - Clinical labels or classes created from the data only
  - MDS-UPDRS scoring
  - Implemented in classes UPDRS and Clinical_UPDRS

Sensor event → data capture → raw data → data ingestion → structured data → data cleaning → signal time series → bio-signal processing → signal features → feature engineering → biomarkers → clustering classification → rating scale score

PDkit pipeline
- Missing and out of range values; data type normalization; indexing; standardised labelling; signal resampling
- Advanced features: gesture verification with DL, data augmentation, signal segmentation, extremity exclusions

Wearables
- Modality: stream processing
- (edge or cloud)
- Format: MQTT. Pub-Sub

Classic unitary and vector instantaneous biomarkers
- \([\text{Power, Amplitude}]^T\)
- Novel longitudinal biomarkers introducing temporal element e.g. same feature calculated at different times
- \(\text{distr}[\text{Power}_1, \text{Power}_2, \ldots, \text{Power}_n]^T\)
- Implemented in class TestResultSet

www.dcs.bbk.ac.uk/~gr
Design objectives:
- Sensitive to patient mobility constraints
- Sensitive to patient cognitive impairments

Approach:
- Constrain user context for reliable interpretation of data
- Encourage frequent use
Test Movements
Unsupervised patient use

- Achieve firm user adherence to the prescribed movements
  - Accept test record only when movement executed correctly
  - Reject test when movement does not match expectations
- Use deep learning to learn movement features
- Apply offline or online (i.e. at the server on in the app)
- Use Tensorflow to learn and apply model
Tremor signal

Original x, y, z features

512 dataset with the artificial feature m

256 dataset with the artificial feature m
Deep Learning Architecture
## Performance using DNN

### Table 1: Accuracy, F1-score, and AUC

<table>
<thead>
<tr>
<th>Classifiers</th>
<th>Accuracy</th>
<th>F1-score</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExtraTrees</td>
<td>0.73</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>BernoulliNB</td>
<td>0.73</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>RandomForest</td>
<td>0.73</td>
<td>0.79</td>
<td>0.83</td>
</tr>
<tr>
<td>GradientBoosting</td>
<td>0.72</td>
<td>0.80</td>
<td>0.83</td>
</tr>
<tr>
<td>Bagging</td>
<td>0.72</td>
<td>0.78</td>
<td>0.83</td>
</tr>
<tr>
<td>AdaBoost</td>
<td>0.66</td>
<td>0.75</td>
<td>0.81</td>
</tr>
<tr>
<td>GaussianNB</td>
<td>0.69</td>
<td>0.75</td>
<td>0.83</td>
</tr>
<tr>
<td>DMLP</td>
<td>0.75</td>
<td>0.81</td>
<td>0.85</td>
</tr>
<tr>
<td>RCNN</td>
<td><strong>0.78</strong></td>
<td><strong>0.82</strong></td>
<td><strong>0.87</strong></td>
</tr>
</tbody>
</table>

### Table 2: Confusion Matrix

<table>
<thead>
<tr>
<th></th>
<th>TP (%)</th>
<th>FN (%)</th>
<th>TN (%)</th>
<th>FP (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExtraTrees</td>
<td>141.52</td>
<td>8.98</td>
<td>13.36</td>
<td>63.14</td>
</tr>
<tr>
<td>BernoulliNB</td>
<td>146.23</td>
<td>4.27</td>
<td>6.92</td>
<td>69.58</td>
</tr>
<tr>
<td>RandomForest</td>
<td>138.39</td>
<td>12.11</td>
<td>16.19</td>
<td>60.31</td>
</tr>
<tr>
<td>GradientBoosting</td>
<td>146.02</td>
<td>4.48</td>
<td>8.12</td>
<td>68.38</td>
</tr>
<tr>
<td>Bagging</td>
<td>135.58</td>
<td>14.92</td>
<td>18.03</td>
<td>58.47</td>
</tr>
<tr>
<td>AdaBoost</td>
<td>128.0</td>
<td>22.5</td>
<td>17.34</td>
<td>59.16</td>
</tr>
<tr>
<td>GaussianNB</td>
<td>116.01</td>
<td>34.49</td>
<td>35.41</td>
<td>41.09</td>
</tr>
<tr>
<td>DMLP</td>
<td>135.73</td>
<td>15.77</td>
<td>28.19</td>
<td>49.31</td>
</tr>
<tr>
<td>RCNN</td>
<td>133.22</td>
<td>18.28</td>
<td>38.38</td>
<td>39.12</td>
</tr>
</tbody>
</table>
• UPDRS exhaustive search of all possible symptoms
• Each patient presents only a few
• Symptoms typically change slowly e.g. 6 months
• ~6 features are predictive of overall score
• Use ML to identify the specific tests that offer the highest inferential power
  – Observer five full tests
  – Apply standard ensemble of randomized decision tree method to rank tests according to predictive strength
  – Select top 3 tests for individualised quick test
1. Digital biomarkers critical for precise disease progression monitoring

2. Google Scholar: 1,000+ papers on Parkinsonian tremor using accelerometers and ML in 2018-19
   – Impossible to replicate and to compare results
   – Differences in data processing and algorithm implementation
   – In most cases, insufficient details provided to replicate algorithm used

3. Common pattern emerging:
   – Machine Learning processing pipeline
   – From raw data to severity assessment (often using MDS-UPDRS scores)
Example: Tremor processing pipeline

From raw accelerometer data to UPDRS score (0-4 scale)

Rest
Tremor of the Left
Hand

<table>
<thead>
<tr>
<th>Raw data</th>
<th>Digital biomarkers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2 30.2 0.1</td>
<td>25.1 0.95 12.1</td>
</tr>
<tr>
<td>2.1 31.1 0.1</td>
<td></td>
</tr>
<tr>
<td>1.6 37.1 0.2</td>
<td></td>
</tr>
<tr>
<td>1.2 25.5 0.0</td>
<td></td>
</tr>
<tr>
<td>1.3 34.2 0.4</td>
<td></td>
</tr>
<tr>
<td>2.1 29.1 0.3</td>
<td></td>
</tr>
<tr>
<td>... ... ...</td>
<td></td>
</tr>
</tbody>
</table>

Fast Fourier Transform
+ Biomarker extraction

Clustering and Classification

UPDRS: 2

Open Source PDkit for python on github

www.dcs.bbk.ac.uk/~gr
PDkit Processing Pipeline

Smartphone apps
Modality: batch processing
Format: JSON, plain TXT

Wearables
Modality: stream processing (edge or cloud)
Format: MQTT. Pub-Sub

www.dcs.bbk.ac.uk/~gr
PDkit Processing Pipeline

- Data converted to standardized representation based on PANDAS dataframes
- Ingestion is extensible so any data format can be loaded
PDkit Processing Pipeline

- Missing and out of range values; data type normalization; indexing; standardised labelling; signal resampling
- Advanced features: gesture verification with DL, data augmentation, signal segmentation, extremity exclusions
- Feature extraction
- 500+ tremor, bradikynesia, tapping, gait, turning and speech features
- Implemented in the TremorProcessor, FingerTappingProcessor, GaitProcessor, SpeechProcessor classes
Classic unitary and vector instantaneous biomarkers:
\[ \text{[Power, Amplitude]}^T \]

Novel longitudinal biomarkers introducing temporal element e.g. same feature calculated at different times:
\[ \text{distr}[\text{Power}_{t_1}, \text{Power}_{t_2}, \ldots, \text{Power}_{t_n}]^T \]

Implemented in class TestResultSet.
PDkit Processing Pipeline

- Supervised and unsupervised learning for scoring
- Clinical labels or classes created from the data only
- MDS-UPDRS scoring
- Implemented in classes UPDRS and Clinical_UPDRS
CUSSP study

- CUSSP at the UCL Institute of Neurology and Homerton Hospital (UK)
  - Details https://clinicaltrials.gov/ct2/show/NCT02937324
  - Data collection completed in May
  - 74 patients
- 20 lines of PDkit source code specify processing protocol
- 2-3 hours of software development
- Can recreate results in 1 hour on standard laptop
CUSSP Clinical Study

Visit 1 Day -60 to 0 (Screening):
Discuss study, PIS given

Visit 2 Day 0 (Eligibility): Written consent, MOCA, PDQ39, Beck’s administered, baseline demographics recorded, App installed

Visit 3 Day 1-150 (Hospital UPDRS): Official Enrolment, assessments as follows with order randomised:
a) Clinical video UPDRS (OFF/ON)*
b) Smartphone UPDRS (OFF/ON)*

Home monitoring period (order of A, B and C pseudorandomised)
(3 consecutive weeks, starting on Day 1-150)

<table>
<thead>
<tr>
<th>Week</th>
<th>6 days of home monitoring+ (method)</th>
<th>Baseline smartphone OFF/ON test on 7th day* Visit 4+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>(method A)</td>
<td></td>
</tr>
<tr>
<td>Week 2</td>
<td>(method B)</td>
<td></td>
</tr>
<tr>
<td>Week 3</td>
<td>(method C)</td>
<td></td>
</tr>
</tbody>
</table>

Visit 6+
Limitations of current practice

- Minimum Detectable Change (MDC95) \(\sim 12\) (range 0-108)
- Typical annual disease progression 3-4 points
- Idealised response model
- Rapid uphill, slow downhill
- Affected by numerous parameters e.g. mood, social interaction, diet, exercise etc
- One sample has extremely limited value
• Tremor signal is not stationary but is often treated as such (not unreasonable due to measurement limitations)
• Consider tremor to be a random process
• Look at temporal aggregates
• Preliminary results suggest far superior MDC95
What is actually measured?

slow start

fatigue settles in

erratic movements

is this PD or advanced essential tremor?

premature end
How precisely can we detect the onset of turning movements? Healthy subjects turn differently than PwP.
• Move from clinician to automated diagnosis and treatment offers great opportunities to realise patient benefits
• Challenges often relate to having to change methods
• This can be intensified by the greater availability of data
• Stationary to dynamic processes, non-linearity
• Validated evidence is time consuming/expensive to collect
Benchmark Performance
• Marco Luchini
• Stefan Kueppers
• Rajesh Pampapathi

re:technica
• Marco Iannone
• Nikos Fragopanagos
• Joan Saez Pons

audience focus
• Theano Moussouri
• Froso Nomikou

UCL IoN
• Bhatia
• John Rothwell
• Ashwani Jha
• Sebastian Schreglmann
• Elisa

Birkbeck College
• Ioannis Daskalopoulos
• Cosmin Stamate
• George Magoulas
• Jenny Vafeiadou
App demo videos

http://www.updrs.net

PDkit analytics toolkit

https://github.com/pdkit/pdkit

Papers

http://www.dcs.bbk.ac.uk/~gr/pubs.html

CUSSP Study Record

https://clinicaltrials.gov/ct2/show/NCT02937324

cloudUPDRS app on the Play Store

https://play.google.com/store/apps/details?id=uk.ac.bbk.dcs.cloudupdrs