A data-driven take on understanding heterogeneity in Parkinson's disease.

Subtype and Stage Inference (SuStain): Computational model of disease progression + Clustering

Parkinson's disease (PD)
- 2nd most common neurodegenerative disease
- Highly heterogeneous

3 no validated biomarkers of progression for PD
- Needed for clinical trials

We estimate a data-driven signature of PD progression subtypes as sequences of measurable abnormality in the Parkinson's Progression Markers Initiative (PPMI)

Study overview: SuStain + KDE-EBM

- Event-based model (EBM) [1, 2] estimates a probabilistic sequence of cumulative abnormality in a set of 3 features using mixture modelling to quantify abnormality
- KDE-EBM [3], Kernel Density Estimation mixture model
- Methodological novelty here: variable bandwidth KDE
- SuStain [4]: split-and-fit KDE-EBM clusters
- Patient subtype and stage inferred from best fitting (max. likelihood) combination of abnormal features
  - "Stage M" ≠ M out of N markers are abnormal

Methods and Data

Event Based Model (EBM) [1, 2] estimates a probabilistic sequence of biomarker abnormality ➞ fine-grained staging tool

- 9 "imaging plus X" [6] features: clinical, cognitive, DaTscan SPECT striatal binding ratio
- Inclusion criteria: complete data

Data

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Disease duration</th>
<th>Age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HC</td>
<td>175</td>
<td>n/a</td>
<td>61 ± 11</td>
</tr>
<tr>
<td>de novo PD</td>
<td>405</td>
<td>0.6 ± 0.5 years</td>
<td>62 ± 10</td>
</tr>
<tr>
<td>SWEDD</td>
<td>56</td>
<td>0.6 ± 0.7 years</td>
<td>62 ± 10</td>
</tr>
</tbody>
</table>

ppmi-info.org

Results: subtypes of Parkinson's

KDE Mixture Models for p(Event) and Cross-Validation


Conclusions

- Parkinson’s disease understanding: fine-grained sequences of subtypes
- New, flexible KDE mixture modelling

Considerations and future work

Method:
- Improve model parsimony

Application:
- Validate in prodromal cohorts
- Other markers: MRI, Vison