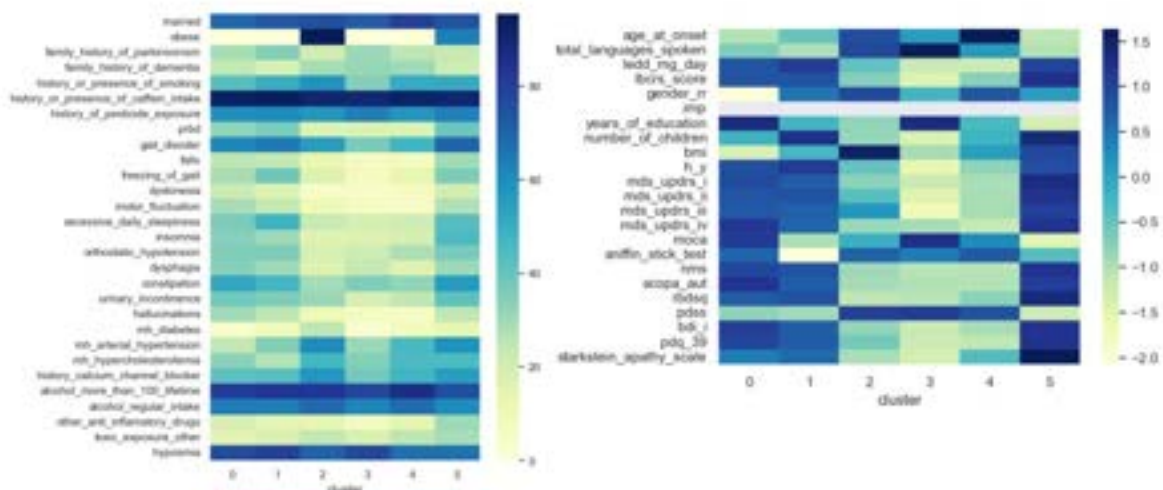


DATA-DRIVEN SUBTYPES OF PARKINSON'S DISEASE USING MACHINE LEARNING IN LUXEMBOURG PARKINSON STUDY

Rawal R; Shoab, M; Pavelka L, Ghosh S ; Landoulsi Z (5) , Pachcek S (1,5); Rauchenberger A (4); Jubal E; Vaillant M,; Aguayo G; Roomp K; Gomes C (3) ; Marques T (3) ; Claire P; Pauly Laure, McIntyre D (1) ; Terwindt O (2,3); Sandt E; Hanff AM; Azaiz K; Perquin M; Ferrand JY; Gantenbein M; May P (5) ; Glaab E (4) , Diederich J N; Satagopam V; Krüger R. (1,2,3) on behalf of the NCER-PD Consortium

- 1) *Transversal Translational Medicine, Luxembourg Institute of Health, L-1445, Luxembourg;*
- 2) *Parkinson Research Clinic, Centre Hospitalier de Luxembourg (CHL), Luxembourg*
- 3) *Translational Neuroscience, Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, L-4362 Luxembourg;*
- 4) *Biomedical Data Science Group, Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-surAlzette, Luxembourg.*
- 5) *Bioinformatics Core, Luxembourg Centre for Systems Biomedicine (LCSB), Esch-sur-Alzette, Luxembourg.*

Parkinson's disease (PD) is characterized by variegated clinical manifestations with different grades of motor and non-motor symptoms, and different biomarkers. With the aim to classify the multifariousness of PD in distinct subtypes and study their distinct clinical phenotypes we used machine learning methods like random forest and K-mean clustering in Luxembourg Parkinson's study. six subtypes were identified first using k-mean clustering. We then used linear regression and random forest to identify significantly different clinical characteristics of each of the subtypes. Visualization and statistical analysis were performed for analyzing the obtained PD subtypes. As a result, 704 patients with idiopathic PD from the baseline data were investigated and six subtypes were identified with 107,108,111,129,132,117 PD patients respectively. The distinct and statistically significant clinical phenotype differences across these six subtypes are summarized in visual representation through heat plots below.



Further work to elucidate the distinctiveness of the subtypes and looking at difference in molecular signatures across different omics platforms along with refining the subtypes is under future pipeline