Characterising the spatiotemporal heterogeneity of neurodegenerative diseases using subtype and stage inference

Multi-omics and big data analytics: From understanding disease heterogeneity to precision diagnostics

Alexandra L Young, Martina Bocchetta, James Cole, Steven CR Williams, Jonathan D Rohrer, Daniel C Alexander

Background
Neurodegenerative diseases are highly heterogeneous, consisting of multiple disease subtypes with different spatiotemporal patterns of pathology. These patterns evolve dynamically with disease stage, making it challenging to disentangle disease subtypes in vivo from biomarker data where precise staging information is unavailable.
Method
Subtype and Stage Inference (SuStaIn) is an unsupervised learning technique that disentangles biomarker heterogeneity into subgroups of individuals (subtypes) with distinct progression patterns (stages). SuStaIn evaluates the optimal grouping of individuals into disease subtypes, where each subtype consists of a sequence in which biomarkers transition between different z-scores. We used SuStaIn in a number of conditions to identify subgroups of individuals with distinct patterns of brain volume loss using structural MRI data. We present results from several datasets, reviewing results in Alzheimer's disease (ADNI dataset), and presenting new results in genetic frontotemporal dementia (GENFI dataset), and in the aging population (UK Biobank). In each dataset we explored genetic, neuropsychological, clinical and biomarker associations with each of the subtypes identified by SuStaIn. We used longitudinal data to assess the consistency of the subtypes and stages assigned by the SuStaIn model at follow-up visits.
Result
In Alzheimer's disease we identified multiple spatiotemporal patterns of neurodegeneration associated with differences in memory and executive function scores. We found a genetic link between a 'limbic-predominant' atrophy pattern and type 2 diabetes through a GWAS and analysis of polygenic risk scores. In genetic frontotemporal dementia we identified and confirmed two distinct atrophy patterns amongst *C9orf72* mutation carriers (N=182). In *MAPT* mutation carriers (N=82), we identified two atrophy subtypes that were strongly associated with different *MAPT* mutations: a 'temporal' subtype, which had a one-to-one mapping with IVS10+16 and R406W mutations, and a 'frontotemporal' subtype, which had a near one-to-one mapping with P301L mutations. The subtype assignments showed strong stability at follow-up visits (99% consistency). In the UK Biobank (N=21390), we identified multiple patterns of brain volume loss associated with distinct characteristics.
Conclusion
Our results demonstrate the utility of SuStaIn for identifying disease subgroups and associating imaging patterns with genetics and cognition. We show that SuStaIn can provide enhanced patient stratification capability across multiple conditions.