Project Title: Systems Medicine of Mitochondrial and biochemical Parkinson's Disease and other related movement disorders (SysMedPD)

Statistical Analysis Plan

Version 1.1

04/10/2017

Sample size

As this part of a multi-site European project statistical review was conducted before submission for EU funding by the lead site (University of Luxembourg). A prospective analysis of the power of the univariate Wilcoxon-Mann-Whitney test was performed. With a total sample size of 120, the power to detect an effect size of 0.5 at a significance level of 0.05 is 0.8 (80%). As such a sample size of 160 should be sufficient to detect meaningful differences as well as have some contingency for patient drop out and anomalous/outlier results that need to be excluded from analysis.

Statistical analysis

1. Comprehensive phenotyping data for the monogenic PD cohort with age and sex matched controls will be used to define a phenotypic signature of monogenic PD with detectable a specific biochemical signature e.g. mitochondrial dysfunction. Phenotyping data will include measurements of metabolites in patient biosamples, biochemical assays of mitochondrial function and clinical data. Standard statistical methods will be used to select the most informative phenotypic assays. Statistical analysis will include univariate and multivariate statistics such as principal component analysis, linear mixed models, partial least squares regression, ANOVA simultaneous component analysis, Lasso shrinkage, and selection method for linear regression.

2. The most informative of the above phenotypic assays will be used to define a biochemical signature for mitochondrial dysfunction in PD. This signature will be used to then stratify idiopathic PD patients by their degree of mitochondrial dysfunction. Those patients who lie at the 'ends of the spectrum' of these pathways (least and most affected) will then have advanced cell models (such as IPSC derived dopaminergic neurons) generated from skin biopsy samples to further study biochemical pathways affected in PD. In the first instance those participants who lie 1 standard deviation outside the control mean of the most informative assays will be considered as those that lie at the 'ends of the spectrum'.