# Enhancing Neuro Imaging Genetics through Meta-Analysis Consortium (ENIGMA)
# Parkinson's Disease Secondary Analysis Proposal

Please complete all fields and return this form by e-mail to:

Ysbrand van der Werf, Max Laansma, Conor Owens-Walton, Eva van Heese, Emile d'Angremont

**Email:** yd.vanderwerf@amsterdamumc.nl; m.laansma@amsterdamumc.nl; conor.owens-walton@loni.usc.edu; e.vanheese@amsterdamumc.nl; e.dangremont@amsterdamumc.nl

## 1. Policy

Members of the ENIGMA Consortium include investigators from different centers around the world who are actively engaged in neuroimaging research and who have contributed results from primary analyses of imaging, genetic data, and/or algorithm development for the purpose of meta-analysis, replication, and/or algorithm testing in a collaborative manner.

Although the data contributed to the ENIGMA consortium consist of group-level summaries and post-estimation statistics rather than raw genotype and phenotype data, there is theoretically a minute risk of determining whether a given individual participated in a study. While the re-identification of samples requires access to the raw genotype data of the target individual and constitutes scientific misconduct, most groups have opted to appoint a gate-keeper approach rather than allowing full public access to the results of their analyses or meta-analyses. Within the ENIGMA-PD working group any consortium member wishing to access the results of specific analyses or meta-analytic results will be asked to complete a short proposal describing why they wish to access the results files from each group, and submit that for review.

All consortium members are encouraged to submit such proposals, to follow up on ideas which the group as a whole cannot pursue, which involve novel analyses, or subsets of the available sites. The ENIGMA-PD working group will screen PD -relevant proposals for scientific interest, and will help enlist members who might be interested in collaborating. Proposals will be discussed on ENIGMA-PD working group calls and emails to encourage the broadest participation.

The proposal will then be posted on an ENIGMA forum page and an email will be sent to all consortium members alerting them to the posting. ENIGMA members will have 14 days from the time of the posting to opt-out of the analysis, ask for clarification, voice concerns or objections and/or give feedback to the proposal. No site data will be shared without the consent of the PI of that site, who may opt to impose specific conditions or limitations on the use of the data; also ENIGMA PIs and members are not required to take part in any proposed project, they can opt out.

If the author of the proposal agrees to the authorship and publication policies of the consortium the access request will be granted to the results files for those groups who have not opted-out of the analysis and a member of the Enigma PD working group or Enigma support group will be assigned as a project liaison. The Enigma support group liaison will be responsible for providing the data and answering any queries relating to the project, and providing the contributing site PIs with updates. If there is no possibility of determining if a particular individual participated in a study (e.g. limited imaging or genetic markers are requested), results from these markers may be sent by the liaison to other sites if available. If genome-wide results are requested from individual groups, the person submitting the proposal may be granted an account on Imaging Genetics Center (IGC) servers or may visit IGC, if desired, to make it easier to complete the analysis. All approved proposals are welcome to use services at IGC. The data can be housed in IGC and will not be transferred or mirrored to other sites.

We request that the 'ENIGMA Consortium' or the specific working group(s), and the liaison person will be listed as co-authors. The ENIGMA Consortium on the byline, or the ENIGMA Working Group on the byline, will reference the PIs of each study, in addition to contributors at their site. In this way the authors contributing data to the consortium will be appropriately acknowledged on any publication.

## 2. Requestor Information

**Date of Submission: 9th September 2024**

**Name: Thomas Welton**

**Institution/Affiliation: National Neuroscience Institute, Singapore**

**Email: thomas_welton@nni.com.sg**

**Have you signed and returned the ENIGMA Memorandum of Understanding? If not, please find the Memorandum of Understanding here.**

## 3. Study proposal

**Proposal title: Comparison of brain morphology among monogenic and sporadic forms of Parkinson's disease**

**Co-author names and e-mail addresses:**

Thomas Welton, Thomas_welton@nni.com.sg

Pablo Mir, pmir@us.es

Start project: January – February 2025
Data inventorisation and organization: January – May 2025
Data analysis: June – September 2025
First draft paper: October 2025

**Proposed Timeline for Completion of Study:**

**Please confirm that you have reviewed the ENIGMA website for potential areas of overlap. If you see a project that may overlap, please list along with any plans for addressing this:**

Confirmed

**Please list any conflicts of interest:**

None

**Please describe the proposed analyses. Include hypothesis, specific results requested, a brief analysis plan and methods, and references.**

Background

Approximately 15% of Parkinson's disease (PD) cases are monogenic, which means they are caused by a mutation in a single gene. The different monogenic forms of PD vary in clinical phenotype and pathological findings. However, the differences in brain morphology are not well described. A better understanding of the differences in brain morphology among monogenic and sporadic forms of PD could lead to improved subtyping and risk modeling. This study aims to compare brain morphology between monogenic and sporadic forms of PD in data gathered from the ENIGMA-PD working group.

Objective

To compare brain morphology between monogenic and sporadic forms of PD.

Data request

*Regional FreeSurfer processed metrics (continuous) from the 7.3.2 version. These metrics will include cortical thickness, cortical surface area, and subcortical volumes.

*Genetic variables (categorical), which are either sporadic (not monogenic) or monogenic including GBA1, LRRK2, SNCA, PINK1, and PRKN. Genetic data of different variants will be included upon availability. We will inventorise data from each site which genetic variables are available and pooled those with paired MRI data.

*Basic demographic and clinical data including age, gender, disease duration, H&Y stage, and UPDRS Part III.

Data processing and harmonization

Quality control: statistical outlier detection and visual inspection of histograms and images based on the output from the FreeSurfer Quality Control (FS-QC) pipeline.

Harmonisation: ComBat batch adjustment for systematic differences in MRI acquisitions across sites.

MRI outcome measures

FreeSurfer 7.3.2.-derived cortical thickness and surface areas for cortical regions and volumetric measures for subcortical regions. The study will incorporate 68 cortical regions of interest (ROIs), 19 subcortical ROIs, along with 22 additional subdivision ROIs based on sub-segmentations across 5 subcortical areas: thalamus, amygdala, hippocampus, brainstem, and hypothalamus.

Statistical models

All analyses will be conducted at two levels. First, a broad analysis will compare the two main genetic groups: monogenic PD and sporadic PD. Second, to address genetic variability within the monogenic PD group, subgroup analyses will be performed after stratifying patients into the five specific genetic profiles: GBA1, LRRK2, SNCA, PINK1, and PRKN.

We will use univariate analyses to test differences in cortical thickness and subcortical volumes across the 87 brain regions of interest (ROIs). For that, we will conduct analysis of covariance (ANCOVA) for each ROI, adjusting for age, sex, education, ethnicity, and disease duration. For the second subgroup analyses of specific monogenic profiles plus the sporadic PD group, further pairwise post-hoc tests will compare individual groups, with corrections for multiple comparisons applied using false discovery rate (FDR) method. Complementarily, we will perform multivariate analyses using random forest to identify complex combinations of regional cortical thickness and subcortical volumes that optimize classification between genetic groups. Random forest will enable the detection of non-linear interactions between the highly dimensional morphological metrics and rank the most important brain regions for group classification. This approach will provide a more comprehensive understanding of the atrophy patterns associated with each genetic group, helping to distinguish between them. Similar to the univariate analyses, we will train two

random forest models: the first to classify the two primary genetic groups (monogenic vs. sporadic PD), and the second to stratify the monogenic PD group into individual genetic forms (GBA1, LRRK2, SNCA, PINK1, and PRKN). In both random forest models, we will include age, sex, education, ethnicity, and disease duration as additional features to account for their potential influence on classification.