medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

AETIONOMY, a Cross-Sectional Study Aimed at validating a new taxonomy of Neurodegenerative Diseases: Study design and subject characteristics

Jean Christophe Corvol,[1] Sarah Bujac,[2] Stephanie Carvalho,[1] Bethan Clarke,[2] Jacqueline Marovac,[2] Graziella Mangone,[1] Olivier Rascol,[3] Wassilios G Meissner,[4,5] Eloi Magnin,[6] Alexandra Foubert-Samier,[4] Hélène Catala,[3] Ioanna Markaki,[8] Panagiota Tsitsi,[8] Raquel Sanchez-Valle,[9] Michael T. Heneka,[11] Jose-Luis Molinuevo,[9] Ullrich Wuellner,[11] Per Svenningsson,[8] Phil Scordis,[2] Martin Hofmann-Apitius,[12] on behalf of the AETIONOMY Clinical Consortium.

Affiliations

1. Sorbonne Université, Assistance Publique Hôpitaux de Paris, INSERM, CNRS, Institut du Cerveau et de la Moelle, UMRS 1127, Department of Neurology, Hôpital Pitié-Salpêtrière, Paris, France
2. UCB Pharma SA, Belgium
3. Centre d'Investigation Clinique, INSERM 1436, Centre AMS du CHU de Toulouse, Centre Expert Parkinson, Service de pharmacologie, F-CRIN, Centre d'Excellence Maladies Neurodégénératives Neurotoul, CHU de Toulouse, Toulouse, France
4. Service de Neurologie, Centre Expert Parkinson, CHU Bordeaux, 33000 Bordeaux, France and Univ. de Bordeaux, Institut des Maladies Neurodégénératives, UMR 5293, 33000 Bordeaux, France
5. Dept. Medicine, University of Otago, Christchurch, and New Zealand Brain Research Institute, Christchurch, New Zealand
6. Regional Parkinson Expert Center and Regional Memory Center (CMRR), Department of Neurology, CHU Besançon, France
7. Dept. of Neurodegenerative Disease and Geriatric Psychiatry, University of Bonn, Bonn Germany
8. Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
9. Alzheimer's disease and other cognitive disorders. Hospital Clínic. Institut d'Investiagació Biomèdica August Pi I Sunyer, Barcelona, Spain.
10. Barcelonabeta Brain Research Center, Pasqual Maragall Foundation, Barcelona, Spain
11. University of Bonn Medical Center, Department of Neurology, German Center for Neurodegenerative disease, Bonn, Germany
12. Fraunhofer Institute for Algorithms and Scientific Computing SCAI, Germany

Corresponding author:
Jean-Christophe Corvol
Clinical Research Center for Neurosciences
Institut du Cerveau et de la Moelle
Hôpital Pitié-Salpêtrière
47/83 Bd de l'Hôpital
75013 Paris, France
jean-christophe.corvol@aphp.fr

Funding: The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under AETIONOMY grant agreement n°115568, resources of which are composed of financial contribution from the European Union's

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution.

Keywords: Neurodegenerative diseases (NDDs), Parkinson's Disease (PD), Alzheimer's Disease (AD), cross-sectional cohort, molecular disease classifications, personalised medicine.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Abstract

Background: Although advances in the understanding of neurodegenerative diseases (NDDs) have led to improvements in classification and diagnosis and most importantly to new therapies, the unmet medical needs remain significant due to high treatment failure rates. The AETIONOMY project funded by the Innovative Medicine Initiative (IMI) aims at using multi-OMICs and bioinformatics to identify new classifications for NDDs based on common molecular pathophysiological mechanisms in view of improving the availability of personalised treatments.

Objectives: The purpose of the AETIONOMY cross-sectional study is to validate novel patient classification criteria provided by these tools.

Methods: This was a European multi centre, cross-sectional, clinical study conducted at 6 sites in 3 countries. Standardised clinical data, biosamples from peripheral blood, cerebrospinal fluid, skin biopsies, and data from a multi-OMICs approach were collected in patients suffering from Alzheimer's and Parkinson's disease, as well as healthy controls.

Results: From September 2015 to December 2017 a total of 421 participants were recruited including 95 Healthy Controls. Nearly 1,500 biological samples were collected. The study achieved its objective with respect to Parkinson's disease (PD) recruitment, however it was unable to recruit many new Alzheimer Disease (AD) patients. Overall, data from 413 evaluable subjects (405 PD and 8 AD) are available for analysis. PD patients and controls were well matched with respect to age (mean 63.4 years), however, close gender matching was not achieved. Approximately half of all PD patients and one At-Risk subject were taking dopamine agonists; rates of Levodopa usage were slightly higher (~60%). Median MDS-UPDRS Part III Scores (OFF state) ranged from 45 (SD 18) in those with Genetic PD to 2 (SD 3) in Healthy Controls. The standardised methodologies applied resulted in a high-quality database with very few missing data.

Conclusion: This is one of the collaborative multi-OMICs studies in individuals suffering from PD and AD involving a control group. It is expected that the integration of data will provide new biomarker-led descriptions of clusters of patient subgroups.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Introduction

Dementia, mainly represented by Alzheimer's disease (AD), affects 44 million people globally, and that figure is set to rise to 135 million by 2050, mostly due to the ageing of the population.[1] Meanwhile an estimated 4-6 million people globally suffer from Parkinson's disease (PD), the second cause of neurodegenerative disease.[2] There is no cure for these devastating diseases, and caring for patients as their disease progresses represents an immense burden for family members, carers, and health and social care systems. The way these diseases are classified is hampering efforts to develop effective, targeted treatments. Currently, diseases are defined largely on the basis of the patient's symptoms and where they occur in the body. There is growing evidence that while two patients may be classified as having the same disease, the genetic or molecular causes of their symptoms may be very different. In other cases, diseases that are currently defined as separate conditions may share a common molecular basis. There is therefore now broad recognition that the way diseases are classified needs to change, and the field of neurodegenerative diseases in particular is considered to be ripe for a rethink.[3]

AETIONOMY was a European project funded by the Innovative Medicine Initiative which intended to develop tools to better classify patients into subgroups based on their underlying pathogenic mechanisms (http://www.aetionomy.eu). During the lifetime of the project the team explored the classification of neurodegenerative disease by dissecting molecular causes of disease and exploring links to clinical evidence. An important activity for the project was the evaluation, and ultimate organisation, of historical data from a range of sources pertaining to the field. This knowledgebase formed the genesis of mechanistic hypotheses that, alongside canonical theories for the drivers of neurodegeneration, were evaluated utilising computational[4] and conventional laboratory methodologies. To facilitate these objectives, a prospective cross-sectional study was initiated that aimed to support the clinical validation of the identified mechanisms in both AD and PD patients. The present publication provides details of the study design, the patient population and their baseline clinical characteristics. It also highlights the study management strategy and the standardised methodologies implemented to ensure quality of the data and sample collection.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Material and methods

Study design

This is a European multi-centre, cross-sectional clinical study, aimed at collecting clinical, brain imaging and biological data on AD and PD patients, individuals at risk of AD and PD, and healthy controls. The primary objective of the study was to validate the new taxonomy of neurodegenerative diseases proposed by the AETIONOMY Consortium in a real cohort of patients, representative of the continuum of AD and PD. Secondary objectives were to describe the correlations between clinical features and biomarkers across neurodegenerative diseases.

Participants

Inclusion and exclusion criteria for the two groups of patients (AD and PD) and their matching controls are described in detail in **Table S1**. For the PD group, subjects were: patients with a diagnosis of idiopathic PD according to UK PD Society Brain Brank,[5] and a disease duration less than 10 years at inclusion; patients with autosomal dominant (*SNCA, LRRK2* or *GBA*) or autosomal recessive (*PARK2*) forms of PD; a group of subjects "at risk of PD" defined as subjects without symptom of PD but either being a first degree relatives of patients with an autosomal dominant form of PD, or subjects with polysomnography-confirmed idiopathic REM sleep behavioural disorders; healthy subjects matched for age and gender without symptoms of PD or other neurological disorders were recruited as controls. For the AD group, we initially aimed to recruit subjects with biologically confirmed AD at the clinical stage, at the prodromal stage, and matched healthy controls. However, we decided to stop enrolment of the AD group shortly after the recruitment started because another IMI-funded project was planning to build a prodromal AD cohort (European Prevention of Alzheimer's Dementia (EPAD) Longitudinal Cohort Study (LCS), http://ep-ad.org/). The AETIONOMY study thus finally mainly focused on the PD group.

Study procedures

Subjects were recruited in six University Hospitals in France (Pitié-Salpêtrière (Paris), Toulouse, Bordeaux, and Besançon), Germany (Bonn), and Sweden (Karolinska University Hospital, Stockholm). Participants were invited to attend a single study visit at the hospital, during which standardised clinical assessments and biological sampling were conducted.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

After obtaining informed consent and the confirmation of inclusion/exclusion criteria, clinical data were collected by a neurologist including: demographics (e.g. age, sex, ethnicity), vital exams (e.g. blood pressure, electrocardiogram), personal and familial medical history, and environmental factor exposure history. Additional disease-specific clinical and neuropsychological assessments were also performed. For the PD group (including controls) these were: Movement Disorder Society-Unified PD Rating Scale (MDS-UPDRS), Hoehn & Yahr stage, Schwab & England Activities of Daily Living (SE-ADL) scale, Non-Motor Symptoms Scale (NMSS), Hospital Anxiety and Depression Scale (HAD), 39-item Parkinson's Disease Questionnaire (PDQ-39), REM Sleep Behaviour Disorder Questionnaire – Hong Kong (HK-RBD), Epworth Sleepiness Scale (ESS), University of Pennsylvania Smell Identification Test (UPSIT) and Stand-walk-sit test. For the AD group (including controls) these were: The 30-item Geriatric Depression Scale (GDS), State-Trait Anxiety Inventory (STAI), Pittsburgh Sleep Quality Index (PSQI), Brain MRI, FCSRT-IR (Buschke) and Clinical Dementia Rating (CDR). The Montreal Cognitive Assessment (MOCA), Mini Mental State Examination (MMSE), Frontal Assessment Battery (FAB), Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) and environment questionnaire including IPAQ were administered to subjects from both groups. Neuropsychological assessments were performed during the visit by trained neuropsychologists. Clinical data were collected in an electronic case report form (eCRF) using the REDCap™ system.[6] The eCRF was developed in collaboration with the Bioinformatics Core Facility of the Centre for Systems Biomedicine (LCSB, University of Luxembourg, UL). All data and analyses produced are stored in a project TranSMART database used and hosted by LCSB, UL and managed by the ICM.

Biological samples

Blood samples were collected in the morning, in a fasted state, into: Serum Separating Tubes (SST, serum collection) for serum, Lithium Heparin tubes (LH) for plasma, and EDTA tubes for DNA extraction. Serum and plasma were frozen and stored at -80°C for biomarker analysis (proteomics, metabolomics, lipidomics, genomics, epigenomics). EDTA plasma samples were also snap-frozen in liquid nitrogen (N2) and stored in N2 for the analysis of neuroinflammation biomarkers. Cerebrospinal Fluid was collected by lumbar puncture. Samples were centrifuged at 4°C within 1 hour, for 10 min at 2,000g and then progressively frozen and stored at -80°C also for biomarker analysis (proteomics, metabolomics, lipidomics). A subset of samples was snap-frozen immediately (within 60 seconds) and stored in liquid N2 for the analysis of neuroinflammation biomarkers. Skin biopsies were also

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

proposed for fibroblast preparations frozen and stored in 10% dimethyl sulfoxide (DMSO) medium. Lumbar puncture and skin biopsy were optional for the participants.

Samples were collected, processed and stored locally in each site following the established standard operating procedures. Samples were then distributed in different centres of the Consortium for specific analyses: genomic analysis (ICM), epigenetics (UKB), proteomics (KI), neuroinflammatory markers (UKB), lipidomics (UKB), AD markers (IDIBAPS), and markers of autophagy and insulin resistance (SARD). Sample management procedures were established prior to the start of the project and clinical centres were trained to ensure standardised collection, preparation and transportation of the samples across countries. All the remaining samples were finally centralised at the ICM biobank (Banque ADN et Cellules – ICM).

Regulatory and ethical approvals

The sponsor of the study was the French institution *Institut National de la Santé et de la Recherche Médicale* (INSERM). We took advantage of the European Clinical Research Infrastructure Network (ECRIN) which took care of all regulatory approvals and applications, under the responsibility of the sponsor, in other European countries, i.e. in Germany and Sweden. The study protocol was approved by the Ethical Review Boards (ERB) of the three participating countries. The study was conducted in accordance with the standards set by International Conference on Harmonization and Good Clinical Practice (ICH-GCP), and to the ethical principles that have their origin in the Declaration of Helsinki (2013). The protection of the confidentiality of records that could identify the included subjects is ensured as defined by the EU Directive 2001/20/EC and the applicable national and international requirements relating to data protection in each participating country.

Sample size calculation

The study was originally designed to recruit approximately 655 subjects in two disease groups (PD and AD). This initial sample size was chosen to allow the identification of subgroups with adequate precision, in accordance with our primary criteria. The primary objective aimed at identifying subgroups of subjects that correspond to a hypothetic causative biological pathway and corresponding to a particular biomarker. Depending on the hypothesis to be validated, the size of the corresponding subgroup may be highly variable. For example, a subgroup defined by genetics could represent only 2% of the total sample. By contrast, a

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

subgroup defined by non-omics markers may consist of 30% to 40% of the total population. The sample size was calculated based upon the idea that we identify a subgroup in X% of the population and we want to measure X% with a pre-defined precision (supplementary material, **Table S2**). Sample-size was estimated using standard software (nQuery. Sample Size and Power Calculation. "Statsols" (Statistical Solutions Ltd), Cork, Ireland) based upon the length of a 95% confidence interval (CI) around the size of the sub-group being "shorter" or "longer". The shorter confidence interval means a more precise estimate and requires more patients. For example, a subgroup consisting of 2.5% of the sample, such as may be found for a subgroup defined by genetics, will have a "reasonable" 95% confidence limit of (1.3%, 3.7%) when 655 subjects are studied.

Statistical analysis

Key demographic, clinical, and disease severity variables are summarised in a descriptive manner only. The intention is simply to provide an overview of the key information that have been collected for each subgroup and to point out some of the main features of this data. Statistical modelling is beyond the scope of this particular manuscript and formal statistical comparisons of the groups are not included. Summary statistics of subject data at entry were produced for each PD subgroup using PC SAS 9.4 (SAS Institute Inc., Cary, NC, USA) and the statistical tools within tranSMART. Due to the low subject numbers, summaries are not provided for the AD patients, nor their controls. Continuous variables are described using means and standard deviations or medians with interquartile ranges where non-gaussian distributions were suspected. Numbers and percentages are provided for each categorical variable. Unless specified otherwise, percentage denominators are the number of non-missing values.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Results

Enrolment of participants

A total of 421 subjects were screened for the study. Eight of these subjects were excluded from the final dataset. The reasons for exclusion were: screen failure (n=3), participant withdrawal upon investigator decision (n=2) and subject decision (n=3) (supplementary material, **Figure S1**). One serious Adverse Event was reported that related to a study procedure (a post-lumbar puncture syndrome with a complete recovery). All data from the subject who experienced this SAE was included in the final analysis set. Overall, data from 413 subjects were available for analysis: 405 in the PD group and 8 in the AD group (**Table 1**). The study achieved its overall objective with respect to the PD group recruitment (412 screened vs. 415 planned, including controls), although the numbers of patients recruited with Genetic PD subjects was fewer than desired (25 screened vs 40 planned). The "at risk PD" subgroup included 25 subjects with idiopathic RBD and 14 with a first degree relative of a genetic PD index case. Due to the early termination of AD recruitment, the number of AD subjects is very small (n=8) and data from this group are not summarised in any further detail in this manuscript. In total, nearly 1500 biological samples were collected (**Table 2**). Biomarker analyses of these samples is currently being undertaken and will be presented in subsequent AETIONOMY Consortium manuscripts.

Demographic and environmental characteristics

Demographic and other subject characteristics are summarised in **Table 3**. Data ascertainment was extremely high for all the variables, as can be seen from the percentages in the final column. PD patients and controls were well matched with respect to age, however close gender matching was not achieved, with the Healthy Control group containing a much higher proportion of females (66%) compared to both the PD and the At-Risk groups (maximum 44% female). The majority of subjects in each group were Caucasian/White (range: 76% in Genetic PD to 98% Idiopathic PD).

Amongst the key environmental factors, rates of current smoking (a protective factor for PD) were low (<13%) across all groups, with more than half of the PD patients having never smoked. The proportion of those who consumed coffee was observed to be much lower in the Genetic PD Group (48%) compared to all other groups where the rates were over 80%. As expected, a history of head injury was most common in the Idiopathic PD group.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Clinical characteristics

Clinical characteristics, disease severity scores and medication use are summarised in **Table 4**. As expected, the majority (64%) of subject with Genetic PD and very few Idiopathic PD subjects or controls (6% for each) reported a family history of PD. Over a third (40%) of the At-Risk PD group also reported a family history of PD. Subjects with Genetic PD had a lower average age at onset (45 compared to 61 years) and a much longer average disease duration (144 compared to 28.5 months) compared with Idiopathic PD subjects.

Dopamine agonists were taken by approximately half of all PD patients and one At-Risk subject (for restless leg syndrome); rates of Levodopa usage were slightly higher (~60%). Consistent with a longer disease duration, the total average Levodopa Equivalent Daily Doses was higher in Genetic PD as compared to Idiopathic PD; the difference being accounted for by higher daily dosages of Levodopa in the Genetic PD group.

A consistent pattern was observed across a number of disease severity scales, such as MDS-UPDRS III, SE-ADL, Hoehn and Yahr, HADS Depression Score and NMSS. Genetic PD subjects tended to have the most severe scores, probably reflecting the longer average duration of their disease, compared to Idiopathic PD subjects. The data from these scales also highlighted that, whilst not as severely affected as the PD subjects, those At Risk of PD were generally more severe than Healthy Controls. As an example, average MDS-UPDRS part III scores were 45 (SD 18) in Genetic PD, 30 (SD 15) in Idiopathic PD, 10 (SD 5) in At Risk subjects and 2 (SD 3) in Healthy Controls. Importantly, MDS-UPDRS part III was performed in OFF-state to better represent the severity of the disease in treated patients. Finally, the majority of the Genetic PD (72%) group, but only a few of the Idiopathic PD group (9-14%), experienced Dyskinesia and Motor Fluctuations.

Average cognitive scores (MMSE, MOCA and RBANS) were very similar between groups. Specifically, there was minimal difference between the At-Risk group and Healthy Controls.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Discussion

The potential for the utility of the large amount of data from multiple disparate sources (literature, public, and private databases) to pave the way for a better classification of patients, based on underlying causes instead of symptoms, was the core objective of the AETIONOMY project. Achieving this goal goes far beyond the scope of any single company or university; the key to success can only come from the broad nature of the project consortium, which brings together expertise in neurodegenerative diseases, molecular biology, clinical research, research ethics, data modelling and simulation, data standards, and patient involvement in research.

The AETIONOMY team's focus on systematically capturing and representing knowledge on neurodegenerative diseases in a computable form to and the generation of a large inventory of multiscale (ranging from the molecular level to the clinical level and cognitive readouts) mechanistic hypotheses for AD and PD lays the foundations for clinical verification of these models in an appropriate and clinically relevant cohort of patients. The role of the AETIONOMY cross-sectional study in this context is therefore to support the validation of novel patient classification criteria provided by these approaches.

We present here the design and the main clinical characteristics of a cross-sectional cohort of subject which aim to represent the continuum of PD pathology. We believe that this clinical dataset and its associated biological collection are suitable for an attempt to stratify the disease according to its underlying mechanisms. As compared to cohorts recruiting only PD patients, our population includes idiopathic and genetic forms of PD, early and late stage of the disease, subject at the prodromal stage (idiopathic RBD) or at the preclinical stage (first degree relatives of genetic PD) of the disease, and healthy controls. Only a few initiatives, like the Parkinson's Progression Marker Initiative (PPMI),[7] have collected in a standardised manner clinical data and biological samples from all these subpopulations representing the whole spectrum of the disease.

In our study, clinical assessments were carefully chosen to allow a comprehensive description of the motor and the non-motor features of the disease. Demographic characteristics are very similar to what has been described in previous studies in PD or prodromal PD,[7-11] although the age at onset is younger than in the general population, probably because patients were recruited in expert centres. As expected, at-risk individuals presented average scores in-between PD patients and healthy controls for most of the rating scales.[12] An extensive biological collection is associated with the AETIONOMY clinical dataset including DNA,

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

plasma, and serum for virtually all patients, CSF for 25% and fibroblasts derived from skin biopsies for 39% of the subjects.

State of the art omics experiments and/or specific biomarker analyses are currently ongoing by members of our Consortium in order to achieve the ambitious goal of stratifying PD patients based on their underlying mechanisms. It is particularly difficult for neurodegenerative diseases since no clear biomarkers have been found to be specifically related to these mechanisms. Current and past exposure to environmental factors and concomitant medication have been collected to adjust for potential confounding factors in upcoming biomarker analyses.

Our study has some important limitations that must be highlighted. Our cohort was initially designed to recruit patients across neurodegenerative diseases, including PD and AD patients. We did not reach our goal in terms of recruitment of AD patients because a competitive project IMI funded, the European Prevention of Alzheimer's Dementia (EPAD) study, was initiated soon after AETIONOMY started. The EPAD cohort will provide complementary data to our cohort which mainly focuses on PD. Although subjects enrolled in our study represent the continuum of PD from healthy controls to subject at risk, and PD patients, AETIONOMY is a cross-sectional study. Future longitudinal prospective studies are needed to investigate the progression of the disease, and determine the different trajectories according to the underlying mechanisms. Finally, the sample size of our dataset remains relatively small to perform unbiased analyses that require very large numbers of patients to correct for multiple comparison testing. Findings from our analyses will thus require replication in independent cohorts.

In conclusion, we provide here a new cohort for investigating the complex biology of neurodegenerative diseases, focusing on PD. Beyond the specific objective of AETIONOMY, we believe that this dataset is valuable for other research on biomarkers, patient stratification, and PD mechanisms. Following the concept of open sciences, our Consortium has implemented a process for data sharing, and the data as well as the samples will be made available on request to the scientific community.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Funding

The research leading to these results has received support from the Innovative Medicines Initiative Joint Undertaking under grant agreement no. 115568, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in-kind contribution.

Acknowledgements

This paper is presented on behalf of the AETIONOMY Clinical Consortium. It was developed with input from the AETIONOMY Steering Committee. We thank all the members of each recruiting centre for their dedicated effort, enthusiasm, promptness and care in the recruitment and assessment of the participants in this study.

The members of the AETIONOMY Clinical Consortium include (in alphabetical order of their country affiliations):
Sarah Bujac, UCB Pharma SA, Belgium
Bethan Clarke, UCB Pharma SA, Belgium
Jacqueline Marovac, UCB Pharma SA, Belgium
Phil Scordis, UCB Pharma SA, Belgium
Stephanie Carvalho, Institut du Cerveau et de la Moelle épinière, Paris, France
Jean-Christophe Corvol, Institut du Cerveau et de la Moelle épinière, Paris, France
Bruno Dubois, Institut du Cerveau et de la Moelle épinière, France
Cecile Gaudebout, Institut du Cerveau et de la Moelle épinière, Paris, France
Graziella Mangone, Institut du Cerveau et de la Moelle épinière, Paris, France
Sylvie Forlani, Banque ADN & cellules, Institut du Cerveau et de la Moelle épinière, Paris, France
Ludmila Jornea, Banque ADN & cellules, Institut du Cerveau et de la Moelle épinière, Paris, France
Philippe Martin-Hardy, Banque ADN & cellules, Institut du Cerveau et de la Moelle épinière, Paris, France
Yassaman Ghassab, Banque ADN & cellules, Institut du Cerveau et de la Moelle épinière, Paris, France
Eloi Magnin, CHU Besançon, France
Alexandra Foubert-Samier, IMNc, Hôpital Pellegrin, CHU Bordeaux, France
Brice Laurens, IMNc, Hôpital Pellegrin, CHU Bordeaux, France
Wassilios Meissner, Hopital Pellegrin Bordeaux, France
Umberto Spampinato, Hôpital Pellegrin, CHU Bordeaux, France
Sylvain Vergnet, IMNc, Hôpital Pellegrin, CHU Bordeaux, France
Olivier Rascol, CHU Toulouse, France
Claire Thalamas, CHU Toulouse, France
Monique Galitzky, CHU Toulouse, France
Fabienne Calvas, CHU Toulouse, France
Fabienne Ory-Magne, CHU Toulouse, France
Christine Brefel-Courbon, CHU Toulouse, France
Michael Heneka, Universitaetsklinikum Bonn, Germany

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Pawel Tacik, Universitaetsklinikum Bonn, Germany
Ullrich Wüllner, Universitaetsklinikum Bonn, Germany
Martin Hofmann-Apitius, Fraunhofer Institute for Algorithms and Scientific Computing SCAI, Germany
José Luis Molinuevo, Barcelonabeta Brain Research Center, Spain
Mircea Balasa, Consorci Institut d'Investigacions Biomediques August Pi i Sunyer, Spain
Beatriz Bosch, Consorci Institut d'Investigacions Biomediques August Pi i Sunyer, Spain
Spain Raquel Sánchez-Valle, Consorci Institut d'Investigacions Biomediques August Pi i Sunyer, Spain
Ioanna Markaki, Karolinska Institutet, Stockholm, Sweden
Per Svenningsson, Karolinska Institutet, Stockholm, Sweden
Panagiota Tsitsi, Karolinska Institutet, Stockholm, Sweden

Disclosures

JC Corvol reports grants from IMI (grant agreement n°115568) in relationship with this study; grants from Sanofi outside this work; scientific advisory board or speaker fees from Ever Pharma, Denali, Biogen, Air Liquide, BrainEver, Theranexus, and BMS outside this work.
SB is an employee of UCB Pharma SA
SC has nothing to disclose
BC is an employee of UCB Pharma SA
JM is an employee of UCB Pharma SA
GM has nothing to disclose
OR reports grants from ANR, CHU de Toulouse, France-Parkinson, INSERM-DHOS, MJFox Foundation, PHRC, European Commission (FP7, H2020) outside of this work; consulting fees from AbbVie, Adamas, Acorda, Addex, AlzProtect, Apopharma, Astrazeneca, Bial, Biogen, Britannia, Buckwang, Clevexel, Denali, INC Reasearch, Lundbeck, Lupin, Merck, MundiPharma, Neuratris, Neuroderm, Novartis, ONO Pharma, Osmotica, Oxford Biomedica, Parexel, Pfizer, Prexton Therapeutics, Quintiles, Sanofi, Servier, Sunovion, Théranexus, Takeda, Teva, UCB, Vectura, Watermark Research, XenoPort, XO, Zambon, outside or this work.
WGM reports fees for editorial activities with Springer, has served as advisor for Sanofi, Lundbeck, Biohaven and Affiris, and has received teaching honoraria from UCB and Boehringer Ingelheim outside of this work.
EM has nothing to disclose
AFS has nothing to disclose
HC has nothing to disclose
IM has nothing to disclose
PT has nothing to disclose
RSV reports scientific advisory board fees from Ionis Pharmaceuticals and Wave Life Sciences outside this work.
MTH has nothing to disclose
JLM has nothing to disclose
UW has nothing to disclose
PS reports honorarium from AbbVie, Safoni Genzyme and Shire outside of this work.
PS is an employee of UCB Pharma SA
MHA has nothing to disclose

Table 1: Study populations

	PD group								*AD group*			
AETIONOMY Cross-sectional study subgroups	**Idiopathic PD patients**	**Genetic PD patients**				**At Risk of PD subjects**	**Healthy controls**	Total	**Prodromal AD subjects**	**At Risk of AD subjects**	**Healthy controls**	Total
		PARKIN mutation	GBA mutation	LRRK2 mutation	SNCA mutation							
Subjects, screened n (%)	255 (54.6)	8 (1.9)	6 (1.5)	10 (2.4)	1 (0.2)	39 (9.5)	93 (22.6)	412	2 (22.2)	5 (55.6)	2 (22.2)	9
Subjects, analysed n (%)	251 (62.0)	8 (2.0)	6 (1.5)	10 (2.5)	1 (0.2)	39 (9.6)	90 (22.2)	405	2 (25.0)	4 (50.0)	2 (25.0)	8

The Parkinson's disease (PD) patient population of AETIONOMY Cross-Sectional Study was divided into three subgroups: Idiopathic PD patients, Genetic PD patients and patients At Risk of PD. The Alzheimer's disease (AD) patient population was composed of two subgroups: Prodromal AD subjects and patients At Risk of AD.
Both of the PD and AD groups of patients were matched with healthy control populations.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 2: The AETIONOMY-CS PD biological collection

AETIONOMY Cross-sectional study subgroups	***DNA Samples***	***CSF Samples***	***Plasma Samples***	***Serum Samples***	***Fibroblasts Vials***	***MRI***
Idiopathic PD (251)	245	75	251	251	85	19
Genetic PD (25)	24	3	25	25	15	0
At Risk of PD (39)	38	5	25	25	15	3
Healthy controls (90)	89	16	90	90	45	8
TOTAL samples (%)	396 (98)	99 (25)	391 (97)	391 (97)	160 (39)	30 (7)

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 3: Demographic and other characteristics

	Genetic PD (N = 25)	Idiopathic PD (N = 251)	At Risk of PD (N = 39)	Healthy Controls (N = 90)	Total records in database (% complete) ∞
Female, n (%)	11 (44.0)	80 (31.9)	13 (33.3)	59 (65.6)	405 (100%)
Age (years)*	59 ± 15	64 ± 9	63 ± 11	63 ± 9	405 (100%)
Ethnicity, n (%)					405 (100%)
Caucasian/White	19 (76.0)	245 (97.6)	32 (82.1)	85 (94.4)	
North African/Arabic	4 (16.0)	3 (1.2)	3 (7.7)	1 (1.1)	
Black African	1 (4.0)	2 (0.8)	1 (2.6)	1 (1.1)	
Caribbean African	0	1 (0.4)	1 (2.6)	0	
Asian	1 (4.0)	0	1 (2.6)	2 (2.2)	
Other	0	0	1 (2.6)	1 (1.1)	
Weight (kg)*	67 ± 13	77 ± 16	74 ± 14	72 ± 14	394 (97%)
Smoking, n (%)					405 (100%)
Current	3 (12.0)	17 (6.9)	5 (12.8)	7 (8.0)	
Past	7 (28.0)	101 (40.7)	18 (46.2)	40 (45.5)	
Never	15 (60.0)	130 (52.4)	16 (41.0)	41 (46.6)	
Alcohol Consumption, n (%)					400 (99%)
Everyday/almost everyday	2 (8.0)	48 (19.4)	9 (23.1)	11 (12.5)	
Regularly but not everyday	0	60 (24.2)	8 (20.5)	23 (26.1)	
Occasionally and in moderation	20 (80.0)	110 (44.4)	19 (48.7)	39 (44.3)	
Occasionally and sometimes excessively	1 (4.0)	12 (4.8)	1 (2.6)	2 (2.3)	
Never/almost never	2 (8.0)	18 (7.3)	2 (5.1)	13 (14.8)	
Coffee Consumption, n (%)					400 (99%)
Regularly	12 (48.0)	207 (83.5)	31 (79.5)	74 (84.1)	
Occasionally	6 (24.0)	20 (8.06)	3 (7.7)	2 (2.3)	
Never or almost never	7 (28.0)	21 (8.5)	5 (12.8)	12 (13.6)	
History of head injury, n (%)	2 (8.0)	64 (25.8)	3 (7.7)	12 (13.6)	400 (99%)

* mean ± SD, N: number of patients with assessable values, ∞ percentage of total subjects (i.e. out of 405), % = (n/N)*100 where the denominator N excludes missing records.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Table 4: Clinical and disease characteristics

	Genetic PD (N = 25)	Idiopathic PD (N = 251)	At Risk of PD (N = 39)	Healthy Controls (N = 90)	Total records in database (% complete) [∞]
Family history of PD (1st/2nd degree), n (%)	16 (64.0)	14 (5.6)	15 (39.5)	5 (5.8)	398 (98%)
Age at onset (years)*	45 ± 16	61 ± 9	N/A	N/A	274 (99%)[§]
Disease duration since diagnosis (months)[+]	144 (125)	28.5 (43)	N/A	N/A	270 (98%)[§]
Concomitant Medication, n (%)	23 (92%)	237 (94%)	30 (77%)	48 (55%)	403 (99%)
PD Medication, n (%)	20 (80%)	198 (79%)	1 (3%)	0	405 (100%)
Dopamine agonist, n (%)	13 (52)	127 (51)	1 (3)	0	405 (100%)
Dopamine agonist LEDD (mg) [+]	160 (140)	157 (142)	98 (0)	N/A	
Levodopa, n (%)	15 (60)	143 (57)	0	0	405 (100%)
Levodopa LEDD (mg) [+]	532 (825)	350 (232)	N/A	N/A	
Total LEDD (mg) [+]	944 (429)	496 (394)	98 (0)	0	
MDS-UPDRS Score*					
Part I	13 ± 7	9 ± 5	7 ± 3	3 ± 3	397 (98%)
Part II	15 ± 9	9 ± 5	1 ± 2	1 ± 1	397 (98%)
Part III (OFF State)	45 ± 18	30 ± 15	10 ± 5	2 ± 3	321 (79%)
Part IV	6 ± 4	1 ± 3	0 ± 0	0 ± 0	391 (97%)
Schwab and England Score, n (%)					391 (97%)
≤ 50%	3 (12.5)	2 (0.8)	0	0	
60% or 70%	3 (12.5)	10 (4.0)	0	0	
80% or 90%	17 (70.8)	202 (81.1)	6 (16.2)	3 (3.4)	
100%	1 (4.2)	35 (14.1)	31 (83.8)	85 (96.6)	
Hoehn and Yahr Score, n (%)					391 (97%)
0	0	5 (2.1)	33 (84.6)	83 (98.8)	
1	0	34 (13.9)	0	0	
2	17 (70.8)	156 (63.9)	6 (15.4)	1 (1.2)	
3	2 (8.3)	46 (18.9)	0	0	
4	0	3 (1.2)	0	0	

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

5	5 (20.8)	0	0	0	
Motor Fluctuations, n (%)	18 (72.0)	35 (14.1)	0	0	391 (97%)
Dyskinesia, n (%)	18 (72.0)	23 (9.3)	0	0	391 (97%)
HADS					395 (98%)
Anxiety Score*	8 ± 4	6 ± 4	7 ± 3	5 ± 3	
Depression Score*	5 ± 3	4 ± 3	3 ± 3	2 ± 3	
NMSS total score*	10 ± 5	9 ± 5	8 ± 4	3 ± 3	388 (96%)
MMSE total score *	28 ± 2	28 ± 2	29 ± 1	29 ± 1	389 (96%)
MoCA total score*	26 ± 4	26 ± 3	27 ± 3	27 ± 3	392 (97%)
RBANS total score*	93 ± 20	92 ± 17	102 ± 14	101 ± 16	377 (93%)

* mean ± SD, [+] median (IQR), [§] % for PD subjects only, [∞] percentage of total subjects (i.e. out of 405), % = (n/N)*100 where the denominator N excludes missing records, *100, LEDD: Levodopa Equivalent Daily Dose, HADS: Hospital Anxiety and Depression Scale, MDS-UPDRS: Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale, MMSE: Mini-Mental State Examination, MoCA: Montreal Cognitive Assessment, NMSS: Non-motor Symptoms Scale, RBANS: Repeatable Battery for the Assessment of Neuropsychological Status.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

medRxiv preprint doi: https://doi.org/10.1101/19004804; this version posted September 26, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

References

1. Collaborators GBDD. Global, regional, and national burden of Alzheimer's disease and other dementias, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2019; **18**(1): 88-106.

2. Collaborators GBDPsD. Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol* 2018; **17**(11): 939-53.

3. Hofmann-Apitius M, Alarcon-Riquelme ME, Chamberlain C, McHale D. Towards the taxonomy of human disease. *Nat Rev Drug Discov* 2015; **14**(2): 75-6.

4. Hofmann-Apitius M, Ball G, Gebel S, et al. Bioinformatics Mining and Modeling Methods for the Identification of Disease Mechanisms in Neurodegenerative Disorders. *Int J Mol Sci* 2015; **16**(12): 29179-206.

5. Hughes AJ, Daniel SE, Kilford L, Lees AJ. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry* 1992; **55**(3): 181-4.

6. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009; **42**(2): 377-81.

7. Marek K, Chowdhury S, Siderowf A, et al. The Parkinson's progression markers initiative (PPMI) - establishing a PD biomarker cohort. *Ann Clin Transl Neurol* 2018; **5**(12): 1460-77.

8. Corvol JC, Artaud F, Cormier-Dequaire F, et al. Longitudinal analysis of impulse control disorders in Parkinson disease. *Neurology* 2018; **91**(3): e189-e201.

9. Heinzel S, Lerche S, Maetzler W, Berg D. Global, Yet Incomplete Overview of Cohort Studies in Parkinson's disease. *J Parkinsons Dis* 2017; **7**(3): 423-32.

10. Malek N, Swallow DM, Grosset KA, et al. Tracking Parkinson's: Study Design and Baseline Patient Data. *J Parkinsons Dis* 2015; **5**(4): 947-59.

11. Postuma RB, Iranzo A, Hogl B, et al. Risk factors for neurodegeneration in idiopathic rapid eye movement sleep behavior disorder: a multicenter study. *Ann Neurol* 2015; **77**(5): 830-9.

12. Fereshtehnejad SM, Yao C, Pelletier A, Montplaisir JY, Gagnon JF, Postuma RB. Evolution of prodromal Parkinson's disease and dementia with Lewy bodies: a prospective study. *Brain* 2019.