PD-MitoQUANT

innovative medicines initiative

IMI2 Project 821522 – PD-MitoQUANT

PD-MitoQUANT – A quantitative approach towards the characterization of mitochondrial dysfunction in Parkinson's disease

WP3 – Project Management and Communication

D3.3 Data management and sharing plan (DMSP)

Lead contributor	Jochen Prehn (1 – RCSI) JPrehn@rcsi.ie
Other contributors	Niamh Connolly (1 – RCSI)
	Amanda Tivan (1 – RCSI)
	Olga Corti (2 – ICM)
	Ayse Ulusoy (3 -DZNE)
	Andrey Abramov (5 – UCL)
	Werner Koopman (6 – RUMC)
	Tania Folke (8 - MIMETAS)
	Ronald Melki (9 - CNRS)
	Inbal Ben Eliezer (11 – TEVA)
	Tina Stummann (12 – LUNDBECK)
	Jonathan Melin (13 – UCB)
	David Dexter (14 – PUK)

Due date	31 July 2019
Delivery date	31 July 2019
Deliverable type	ORDP (Open Research Data Pilot)
Dissemination level	Public

Description of Action	**Version**	**Date**
	V1.0	18 Feb 2019

Document History

Version	**Date**	**Description**
V0.0	27 May 2019	First Draft
V0.1	21 Jun 2019	Comments – incorporated comments from the Project Coordinator DPO
V0.2	25 Jul 2019	Comments – incorporated comments from the Project Coordinator
V1.0	26 Jul 2019	Final Version

efpia

Table of contents

Publishable Summary

Mitochondrial dysfunction is implicated in Parkinson's Disease (PD), but detailed understanding of the cause and effect in alpha-synuclein (αSyn) toxicity is lacking. Through provision of quantitative and systematic characterisation of mitochondrial dysfunction, PD-MitoQUANT will provide unprecedented understanding of the role of mitochondrial dysfunction in PD, identify and validate novel disease biomarkers, and propose innovative therapeutic targets that can be further progressed by industry partners within the consortium. The consortium leverages multi-disciplinary expertise in the fields of αSyn biochemistry, iPSC-derived PD models, mitochondrial function and structural analysis, proteotoxicity, ER stress and UPR signaling, systems biology of mitochondrial function, and *in vivo* animal models. The purpose of the Data Management and Sharing Plan is to ensure harmonized data collection throughout the PD-MitoQUANT project, ensuring that all data collected and generated will adhere to **FAIR** (**f**indable, **a**ccessible, **i**nteroperable and **r**eusable) guidelines, whilst also adhering to required EU General Data Protection Regulation (Regulation (EU) 2016/679) (GDPR), and national data protection legislation, contributing to ethical data sharing beyond the life of the project proposal.

Methods

1. Data summary

The overall objective of this project is to provide a systematic and quantitative characterisation of mitochondrial dysfunction in cellular and *in vivo* models of Parkinson's disease (PD), focusing on defined, innovative αSyn toxicity models, using robust protocols and positive controls previously established by members of the PD-MitoQUANT consortium. By incorporating ageing studies, proteostasis studies, computational modelling and data integration approaches, PD-MitoQUANT will identify signatures of progressive mitochondrial dysfunction in PD pathogenesis, and deliver a platform for the identification of novel therapeutic targets for the treatment of PD and other neurodegenerative disorders.

1.1 State the purpose of the data collection/generation:

The purpose of data collection in the PD-MitoQUANT project is to directly address the overall research objective, which is to provide a new and unprecedented understanding of the role of mitochondrial dysfunction in PD. Identification and validation of novel disease signatures and biomarkers will be undertaken with proposed innovative therapeutic targets further evaluated by industrial partners within the PD-MitoQUANT consortium. PD-MitoQUANT data will be scientifically valuable and made **f**indable, **a**ccessible, **i**nteroperable and **r**eusable (**FAIR**), while adhering to required EU General Data Protection Regulation (Regulation (EU) 2016/679) (GDPR), and national data protection legislation, contributing to ethical data sharing beyond the life of the project proposal.

Notably, 95% of all data generated during the course of the PD-MitoQUANT project is non personal research data, however the remaining 5% of data generated will be from RNA sequencing of human and animal biological material. As human genetic data is considered to contain sensitive health and non-health-related information about the individuals and their family members, adopting adequate privacy safeguards is paramount when processing genetic data for research or clinical purposes (Article 4 (1) Article 4 (13) Recital 34). All human sample data used in the PD-MitoQUANT project are pseudonymised; that is, only referred to by an identifier that is not associated with the individual. The researchers processing the data (i.e. PD-MitoQUANT consortium partners) are not themselves in possession of the key of how the identifier relates to the individual.

1.2 Explain the relation to the objectives of the project:

With respect to the data collected and generated during this project, outlined below are specific workflows with regard to how data will be utilized and further expanded.

Data will be obtained from quantitative analysis of key aspects of mitochondrial function and dysfunction in cultured dopaminergic, cortical and/or hippocampal neurons challenged with αSyn toxicity. The data generated will be analyzed to relate defects in mitochondrial function to structural alterations, performing integrated quantitative analyses of mitochondrial morphofunction. Through PD-MitoQUANT, a centralized platform for the use of established human iPSC-derived neuron models from healthy controls and αSyn overexpressing PD patients will be generated. Data pertaining to these models will be available for all consortium members. Data pertaining to human iPSC cells will be pseudonymised at source and all ethics, data protection impact assessments (DPIA) and appropriate GDPR compliant consents will be carried out by the individual partners collecting the data and provided to the Data Coordinator for the consortium, RCSI (Data Manager) (Tasks 4.1, 4.3, 4.4).

In coherence with industry partners, data will be generated from the characterization of the effects of αSyn toxicity on mitochondrial dysfunction in a defined set of standardised, *in vivo* PD disease models as outlined within the Consortium Agreement (CA). PD-MitoQUANT will generate imaging data through delivery of state-of-the-art platforms for super-resolution imaging of alterations in mitochondrial (ultra)structure during αSyn toxicity *ex vivo* and *in vivo*. Employing standardised consortium protocols, PD-MitoQUANT will generate data through the performance of a discrete set of RNA sequencing, proteomics, and stress response and proteostasis studies. These will relate mitochondrial function and structure alterations to proteostasis dysregulation, stress and cell signaling, and identify how αSyn species interact with intracellular and mitochondrial proteins. Research data will be generated through identification of master regulators of αSyn toxicity which will be achieved by dedicated bioinformatics and data integration platforms at the RCSI Centre for Systems Medicine and by the small-medium sized enterprise (SME) partner GeneXplain.

We will also use these data as inputs into experimentally validated, systems biology models of the mitochondrial respiratory chain, carbon metabolism and proteostasis/ER stress signaling. Incorporating this network information will improve our ability to identify early defects and targets for therapeutic intervention, as successfully demonstrated in previous EU consortia. The consortium will employ data-driven statistical modelling approaches to deliver a combinatorial mitochondrial signature of αSyn toxicity, combining single biomarker, network-derived, and morphological signatures and using established built-in quality control (QC) strategies. PD-MitoQUANT will perform mitochondrial function and structure studies on innovative microfluidic and 3D 'Organ-on-a-Chip' (OrganoPlate®) technologies. Data generated from such studies will be shared with other consortium members. The consortium will also generate data through the investigation of identified signatures and explore novel biomarkers in innovative PD models that incorporate the effects of ageing on αSyn toxicity. These models will include the direct cellular reprogramming of aged patient-derived fibroblasts into dopaminergic neurons which incorporate major ageing aspects by maintaining epigenetic marks.

PD-MitoQUANT will establish an effective dissemination and communication, data management and intellectual property (IP) management strategy; which will ensure protection and exploitation of research findings to the ultimate benefit of PD patients. State-of-the-art data quality, transfer and management strategies will be undertaken, ensuring data is FAIR and complies with Clinical Data Interchange Standards Consortium (CDISC) standards for clinical research and CDISC Standards for Exchange of Nonclinical Data (SEND); in addition to a robust ethics management infrastructure, ensuring accessibility and reproducibility of findings.

1.3 Specify the types and formats of data generated/collected:

Non-personal research data generation is the primary data output of the PD-MitoQUANT project. Personal data collection is not a component of the work package task outlines for the PD-MitoQUANT project. However, RNA sequencing data is considered as the generation of personal data and management of such is subject to GDPR regulations and managed accordingly in the PD-MitoQUANT project.

1.4 Types of Data Generated/Collected:

Data will be generated during the PD-MitoQUANT project from experiments conducted at all partner sites in a variety of formats. Such forms of data include:

Non personal data

- Research data generated from experimentation on cultured human induced pluripotent stem cells (iPSCs), primary cultured mouse cortical and/or hippocampal and midbrain cells, as outlined in Table 1.
- In addition, data from *in vivo* animal models, e.g. imaging data, data relating to toxicity studies, morphofunctional study data, 'omics data (transcriptomics, proteomics) and other biomarker data will also be generated as outlined in Table 1.
- Simulation data from computational models.
- Quantitative data: experimental data, medical electronic health records, administrative records, images, collection of data pertaining to partner interaction, consortium meetings, and project co-ordination.
- Derived or compiled data: after data mining or statistical analysis has been done.

Sensitive personal data

- Pseudonymised genetic data from RNA sequencing of human iPSCs and human PD patient tissue.
- Pseudonymised clinical data from patient's records will be provided to the PD-MitoQUANT consortium members from the PUK brain bank to study mitochondrial dysfunction in Parkinson's patient tissue samples in accordance with GDPR regulations.

Over the course of the Project, subsequent versions of the Data Management and Sharing Plan (DMSP) will include more information on the specific datasets generated and any alterations which occur as a result of the UK leaving the European Union.

1.5 Formats of Data Generated/Collected:

Non-personal data generated during the PD-MitoQUANT project can be in many different forms: textual, numerical, databases, geospatial, images, audio-visual recordings, and data generated by machines or instruments. Digital data exists in specific file formats which are coded so that a software programme can read and interpret these data. Using standard and interchangeable or open data formats ensures longer-term usability of data. The format and software in which research data are created usually depend on how researchers choose to collect and analyze data, which is often determined by discipline-specific standards and customs. All appropriate data generated during PD-MitoQUANT will be saved in a format outlined in section 2.3 (Table 3), and personal data will be handled according to GDPR legislation.

1.6 Specify if existing data is being re-used (if any):

During collection of information for the economic evaluation of model/assay development all partners will access literature (published and unpublished) and websites with open access data (e.g. PubMed). In addition, some partners will use previously generated models/data/pipelines (i.e. generated prior to PD-MitoQUANT) to inform and develop their Tasks.

1.7 Specify the origin of the data:

PD-MitoQUANT research activities and statistical analysis of data generated will take place in Ireland, France, Belgium, Italy, Germany, United Kingdom, Denmark, Netherlands and Israel. No work will take place until approval of the appropriate local and/or national ethics committees has been granted, which will include all the necessary and appropriate permissions to access participants and data, ensuring that the research complies with national and EU directives. In this regard, best practice guidelines will be employed and therefore all work conducted will align with the Commission's ethical principles for Horizon 2020. As established in the Consortium

Agreement (CA), each partner agrees to comply with all obligations and requirements of its corresponding national data protection legislation and GDPR requirements; providing all legal documents and certifications, required for compliance with such legislation, to the consortium Data Coordinator (RCSI) for data protection control.

1.7.1 Human

(a) Human induced pluripotent stem cells (iPSCs)

All data which relate to human iPSC donors will be pseudonymised by the Consiglio Nazionale delle Ricerche (CNR) partners prior to its distribution to other partners within the consortium, and the key for decoding will stay with the CNR and will not be made available to other partners. This coding includes not only the removal of personal identifiers and information but also of clinical information that could potentially be used to identify a single individual. Genetic data, used by CNR to confirm genetic criteria of αSyn of human patient donors, will not be provided to the consortium. As such, data generated through experimental use of human iPSCs during the course of the PD-MitoQUANT project will be considered non-personal data. Notably, however, data handling standards, procedures, and best practices will be adhered to at all times. The Data Coordinator (RCSI) will ensure that the appropriate documentation for data protection impact assessments (DPIA), recording of processing activities (local ethics approval) and collection of evidence for obtaining consent (informed consent and patient information leaflets), as completed in WP4 (Task 4.2-4.10), will be available both for inspection by the funding body and to ensure compliance with additionally required legislation which may arise during the course of this project.

Genetic (personal) data generated from the RNA sequencing of human iPSCs will be managed in accordance with appropriate GDPR criteria, and all legal requirements under EU (regulation (EU) 2016/679) under local and national law will be applied by each respective research team prior to personal data processing.

(b) Parkinson's UK (PUK) Repository Tissue Samples

PUK will facilitate access by all partners within the PD-MitoQUANT consortium to the Parkinson's UK Tissue Bank (offering bespoke tissue collection). The PUK Tissue Bank is hosted by Imperial College London (ICL), which is also the legal custodian of the Tissue Bank. Any activities related to the tissue bank and the provision of Parkinson's disease tissue and associated data to other project beneficiaries will be carried out with the support of personnel and resources from Imperial College London. All data which relate to human research participants will be pseudonymised by the PUK Tissue Bank at ICL, prior to its reaching other partners within the consortium, and the key for decoding will stay with the PUK Tissue Bank and will not be made available to others. Genetic data (DNA-sequencing), used by the PUK Tissue Bank at ICL to validate familial Parkinson's cases, will not be provided to the consortium. Coded clinical data will be available (e.g. Parkinson's symptoms, drug treatments, etc.) as a clinical summary from the donor medical files, but in such a manner to prevent subject identification. As such, data associated with human PD tissue will be considered non-personal; however, data handling standards, procedures, and best practices will be adhered to. The Data Coordinator (RCSI) will ensure that the appropriate documentation to confirm completion of data protection impact assessments (DPIA), recording of processing activities (local ethics approval) and collection of evidence for obtaining consent (informed consent and patient information leaflets) will be available for inspection by the funding agency, completed as part of WP4 (Task 4.2-4.10).

Genetic (personal) data generated from the RNA sequencing of human PD tissue will be managed in accordance with appropriate GDPR criteria, and all legal requirements under EU (regulation (EU) 2016/679) under local and national law will be applied by each respective research team prior to personal data processing.

1.7.2 Animal

Data generated through the use of *in vivo* mouse models, and primary cortical and/or hippocampal and midbrain neurons, and *in vivo C. elegans* and *Drosphila* invertebrate models.

In order to investigate and validate molecular targets and biomarkers identified in cultured cells, and further characterize mitochondrial dysfunction *in vivo*, well-established αSyn models will be adapted in coherence with members of the consortium. The data generated from such research will be confined by the same ethical and data regulations as imposed at national and EU level. Data pertaining to ethical approval for generation of *in vivo* models for use within the PD-MitoQUANT project will be provided to the Data Coordinator (RCSI) as completed in WP4, in addition to copies of all import and export licenses for animal models between partners, should this be applicable,.

Animal models which will be used include: a) AAV-mediated human αSyn over-expression in substantia nigra dopaminergic neurons, causing reproducible protein aggregation and dopaminergic neurodegeneration within 6-8 weeks; b) αSyn fibril injection into the striatum, causing αSyn pathology spreading to connected brain areas within a few months, also resulting in progressive nigral dopaminergic cell loss; and c) Validation of effects of candidate master regulators on αSyn-induced changes in mitochondrial function in both *C. elegans* and *Drosophila*. Data resulting from in-depth quantitative investigation of mitochondrial function will be generated from all models. This will include *ex vivo* bioenergetics data, data from mitochondrial morphology studies, and the extensive validation of mitochondrial dysfunction signatures identified *in vitro*. Additional data will be generated through targeted validation of signatures *in vivo,* including genetic models (Parkin- and PINK1- deficient animals, available at The Institut du Cerveau et de la Moelle épinière (ICM-Brain & Spine Institute)) and models of αSyn spreading from the German Center for Neurodegenerative Diseases (DZNE).

1.8 State the expected size of the data (if known):

The format and scale of the datasets generated in this project are summarized in Table 1. At this stage (Version1-initial) of the DMSP, the size of each dataset is estimated. Additional detail of the format and size will be updated during each revision (new version) of the DMSP as more information becomes available. Associated metadata (depending on each datatype) will be collected and stored at the data collection site.

Table 1. Format and size of data sets within PD-MitoQUANT

Data Type	Partner/ Institution	Format	Size
Immunohistochemistry & PLA	RCSI	.xls, High resolution confocal images (.lsm)	<2 GB, 30GB
RNASeq	RCSI	FastQ, .txt file	2GB, <0.2GB/sample
Metabolomics profiling	RCSI	.raw, .xls	200MB/sample
Proteomic profiling (Mass Spec)	RCSI	.xls	<2GB
RPPA	RCSI	TIFF, .xls	20GB
Cell viability Assay	RCSI	.xls, .fcs, .wsp,	20MB
HSC experiments	RCSI	.mdb, .tif, .cp	50GB/experiment
Seahorse Expts	RCSI	.xls, Wave software files: .asyr	2MB/experiment
Confocal imaging	RCSI	.xls, .lsm, TIFF	500MB/experiment
Morphofunction expts	RCSI	.xls, .lsm, TIFF	500MB/experiment
Microfluidics expts	RCSI	.xls, .lsm, TIFF	500MB/experiment
Cell viability Assay	CNRS	.xls, .fcs, .wsp,	20MB
Biochemistry & Cell biology	CNRS	kinetic measurements .xls, Digital transmission electron microscopy images .TIFF gel electrophoresis images and/or western blot images .TIFF intensity quantification .xls; fluorescence microscopy cell imaging .TIFF	30GB
Proteomic profiling	CNRS	Mass Spectrometry output files .mzML	Medium
Immunocytochemistry	MIMETAS	TIFF	2.5GB per Organoplate
Live/dead assay	MIMETAS	TIFF, .txt	2.5GB per Organoplate
Mitochondrial assay - TMRM/MitoSOX/MitoPerOX	MIMETAS	TIFF, .xls	2.5GB per Organoplate
Calcium imaging	MIMETAS	TIFF, AVI, .csv	50GB per Organoplate
Neurite outgrowth assay (calcein-AM)	MIMETAS	TIFF, .xls	2.5GB per Organoplate
Phase contrast imaging	MIMETAS	TIFF	1GB per Organoplate
qPCR	MIMETAS	LC96P, .txt, .xls	500kB per Organoplate
Electrophysiology MEA	Teva	.raw, .spk	100MB per experiment
Electrophysiology patch clamp	Teva	.abf	2MB per experiment
Seahorse Assay	Teva	.asyr, .xls	2MB per experiment
High-content microscopy images	RUMC	.TIFF	771 MB per 96-well plate

Data Type	Partner/ Institution	Format	Size
High-content microscopy movies	RUMC	.SEQ	~3.5 GB per 96-well plate
High-content microscopy data	RUMC	.xls, .OPJ, .M files	<100 MB per 96-well plate
IHC and image analysis (mouse)	ICM	TIFF, .ndpi, .wfml, .mcr, .xls	1GB (maximum)
IHC images (Drosophila)	ICM	.TIFF	0.1-0.5GB/sample
High-content imaging data (cells)	ICM	.CO1, .TIFF, .nd2	1 GB/well plate
Conventional microscopy data (cells)	ICM	.CZI, .TIFF	20GB
Spinning disk microscopy images (cells)	ICM	.TIFF	70 MB/sample
RNASeq	ICM	FastQ, .txt file	2GB, <0.2GB/sample
Metabolomics profiling	ICM	.raw, .xls	200MB/sample
Proteomic profiling	ICM	Mass Spectrometry output files, mzML	Medium
RPPA	ICM	TIFF, .xls	20GB
Cell viability Assay	ICM	.xls, .fcs, .wsp,	20MB
Videos of animal behavior (Drosophila)	ICM	.mov, .avi	1-5GB/sample
Liquid scintillation spectrometry ([3H]-dopamine uptake)	ICM	.xls	< 100 MB
Bioluminescence (import monitoring, cells)	ICM	.txt, .xls	< 100 MB
Respirometry data (tissue)	ICM	.dld, .xls	3MB per experiment
Pseudonymised clinical neuropathology autopsy report data will be included with each PD brain tissue sample.	PUK	.txt	<15kB per sample
Mitochondrial assay - TMRM, NADH, FAD	UCL	TIFF, .xls	30GB
ROS measurements, MitoSOX, DHE, MitotrackerROS	UCL	TIFF, .xls	20GB
Mitochondrial and cytosolic calcium imaging	UCL	TIFF, .xls	20GB
Microscopy images	DZNE	.czi, .tiff	<2 GB per image
Stereological cell counts	DZNE	.DAT, .xls	<100kB per sample
Microscopy images	DZNE	.czi, .tiff	<2 GB per image
FRET assays, cellomics, confocal images	Lundbeck	TIFF, .xls	<50GB

1.9 Methodologies for data collection/generation:

In order to ensure that each member within the consortium partake in harmonized protocols for research and data generation, all consortium members undertake to collect each datatype via in-house (electronic) data capturing systems as listed in Table 2.

Table 2. Methodologies for data collection within PD-MitoQUANT

Data Type	Partner/ Institution	System
RPPA	RCSI	Microvigene software (Vigenetech Inc)
RNA Seq	RCSI	Illumina HiSeq4000.
Mass Spectrometry	RCSI	Liquid Chromatography-Tandem Mass Spectrometry (Nanoflow Ultimate 3000 LC and Q-Exactive mass spectrometer [Thermo])
Confocal imaging and analysis	RCSI	Zeiss LSM confocal, Image J, Metamorph
Seahorse assays	RCSI, Teva	Agilent Seahorse XF Analyzers
High Content Screening	RCSI	Cellomics, Zeiss
Data Analysis	Several partners	Matlab, GraphPad Prism, Excel etc.
Electron microscopy	CNRS	Jeol 1400 electron microscope equipped with a Gatan Orius CCD camera (Gatan)
Fluorescence, absorbance and chemiluminescence measurements	CNRS, Lundbeck	Spectrophotometer (e.g. Agilent, Cary 60 UV-VIS spectrophotometer), Fluoremeter (e.g. Agilent, Cary Eclipse Fluorescence and Bioluminescence (e.g. Chemidoc imaging system (Biorad)); quantification (e.g. Image Lab v.5.2.1. software (Biorad)).
Mass spectrometry	CNRS	Maldi MS (tof-tof on 50800, AB Sciex)
Fluorescent and phase contrast images	MIMETAS	High-content imaging systems (ImageXpress Micro Confocal, ImageXpress Micro XLS), and MetaXpress software
qPCR	MIMETAS	LightCycler 480 Real-Time PCR machine, LightCycler 96 software
Electrophysiology patch clamp	Teva	Patch-clamp system/electrophysiology set-up
High-content microscopy images: TIFF	RUMC	BD Pathway 855 automated microscopy system (Becton Dickinson)
High-content microscopy movies: SEQ	RUMC	Image Pro Plus software (Media Cybernetics)
High-content microscopy data: M	RUMC	MATLAB software (The Mathworks)
High-content microscopy data: XLSX	RUMC	Excel software (Microsoft Corp)
High-content microscopy data: OPJ	RUMC	Origin Pro plus software (Originlabs)
IHC (Drosophila)	ICM	Confocal microsocopy
Videos of animal behavior (Drosophila)	ICM	Standard smart phone camera
IHC microscopy images (mouse)	ICM	NonoZoomer 2.ORS (Hamamatsu)
Immunohistochemistry image analysis (mouse)	ICM	Mercator software (Explora Nova); Calopix software (Trybvn Healthcare)
High-content imaging data (cells)	ICM	Arrayscan XTI (HCStudio); Nikon A1R HD25 (NIS Elements)
Conventional microscopy data (cells)	ICM	Axio Observer7 (Zen)
Spinning disk microscopy images (cells)	ICM	Spinning disk CSU-X1 confocal microscope (Leica); MetaMorph software

Liquid scintillation spectrometry ([3H]-dopamine uptake)	ICM	Packard Tricarb 4910TR
Bioluminescence (cells)	ICM	FDSS 7000 Functional Drug Screening System (Hamamatsu)
High resolution respirometry (tissue)	ICM	Oxygraph-2k with O2k-Fluo LED2-Module (O2k, OROBOROS Instruments); DatLab software
Clinical and neuropathology autopsy reports	PUK	Consultant assessment and diagnostic reports
Live cell fluorescent imaging	UCL	Zeiss LSM software
Microscopy images	DZNE	ZEN software
Image analysis	DZNE	Imaris and Image J
Stereology	DZNE	Stereo investigator

1.10 Outline the data utility: to whom will it be useful?

Data generated during the course of the PD-MitoQUANT project will be useful for the following stakeholders:

- PD-MitoQUANT Partner Organisations (including academic and SME parties)
- External Academic organisations and SMEs
- Pharmaceutical companies
- Clinicians and Health Care providers/payers
- Patient advocacy organisations, people living with the condition
- Charity and funding bodies, regulatory agencies

2. FAIR data

2.1 Making data findable, including provisions for metadata:

All consortium members will take part in the FAIR (**F**indable, **A**ccessible, **I**nteroperable and **R**e-usable) initiative which will harmonize metadata and data formats. This is essential for the PD-MitoQUANT consortium to have the widest impact scientifically, socially and economically. Good documentation and file management from the outset and during the course of a research project makes material understandable, verifiable and reusable by others in accordance with FAIR Data Principles. PD-MitoQUANT has appointed a Data Steward from Coordinator RCSI, who will ensure that PD-MitoQUANT data is properly collected, annotated, archived and preserved into the future to facilitate and simplify the ongoing process of use and reuse. The Data Steward will coordinate with all sites to ensure that appropriate training, consistent data collection and metadata standards are maintained throughout. The interaction between the Data Steward and all partners for the compilation of the DMSP will also ensure that the data is FAIR from the outset.

2.1.1 Making (meta)data findable including internal data sharing:

In order to ensure that data generated from the PD-MitoQUANT project which is published or deposited in a repository is findable, standard identification mechanisms such as a Digital Object Identifier (DOI), Dataset Accession Number (DAN), or GEO accession number (GSE) will be generated. Additionally, the data will be described with rich metadata, which will clearly and explicitly include the identifier of the data it describes. The source in which the (meta) data is indexed will be searchable immediately after publication. For a number of smaller or manually created data sets generated within individual research tasks or work packages, persistent identifiers may not be required. More specific information will be made available on such datasets in future updates of this document.

Internal data sharing: For data that is to be shared amongst the consortium, RCSI will work with the PD-MitoQUANT partner UCB who will provide access to a secure data sharing platform

(Sharepoint) for efficient data transfer between PD-MitoQUANT partners and to provide a data standardization concept to facilitate data queries, transfer and joint analyses.

The RCSI IT research infrastructure, which will support PD-MitoQUANT data storage and will comprise an IT Systems administrator who will work with research teams and the IT Research storage and compute resources, will provide a dedicated network and storage platform for research data. This infrastructure will be available from September 2019 (M7).

The PD-MitoQUANT data sharing protocol (through Sharepoint) will support the sharing of standard data type formats, and facilitate data exchange via data downloading facilities. During the course of the project, ***sharing of embedded or attached research data through email communication is not deemed a secure means of data exchange***. Documents containing data to be shared will be saved in an appropriate format, encrypted, and uploaded to the PD-MitoQUANT Sharepoint-Data Sharing folder with appropriate nomenclature. The code for decryption will be provided separately to the intended partner for access.

To ensure fair usage of the data sharing facility provided by UCB and avoid the use of Sharepoint as a data storage location, the following protocols have been established which also ensures GDPR-compliant personal and non-personal data sharing between partners within the PD-MitoQUANT consortium.

1. Each individual partner is responsible for their own (meta) data collection, storage, back up and security.
2. In the event that hardware used to store PD-MitoQUANT data is lost/stolen, this information has to be detailed to the partners' local Data Protection officer (DPO) and the PD-MitoQUANT Project manager within 48-72 hours.
3. Research Data must be encrypted prior to transfer to other partners.
4. Encrypted research data must be uploaded **for a temporary period (7 days maximum)** to the MitoQUANT PD Sharepoint **Data Sharing** folder. The sending partner must contact the intended recipient through email and include the RCSI Data Manager in correspondence.
5. The recipient partner has a total of 7 days to download **and delete** the file of interest from the Sharepoint platform.
6. If this data has not been removed after a period of **7 days from the initial posting**, then the Data Manager (RCSI) is responsible for removal of this file from its temporary storage location. This routine maintenance of the Sharepoint Data Sharing folder is to ensure efficient and appropriate use of this sharing platform.
7. In the event that data needing to be transferred are in excess of 1GB please do not use the Sharepoint facility to share data. Please contact the RCSI Data Manager who will coordinate large file transfer with the support of Dr. Jonathan Melin, UCB (Sharepoint Host Coordinator)

2.1.2 Naming Conventions and considerations:

For data saved relating to the PD-MitoQUANT project, wherever there are standard naming conventions required, '-PDMQ' should be used following such conventions. In advance of data being made accessible, optimal keywords will be determined which are not restrictive; and, where necessary, version numbers will be provided. With respect to data sharing and collaborative edits, partners should maintain a key/record to all shared data documentation versions which can be included in each document or in the metadata associated with each document, for example:

File Name/Version	Responsible	Changes to file	Last Amended
Image3_1.0-PDMQ	RCSI_AA	Original document	01.01.2019
Image3_1.1-PDMQ	DZNE_AB	Minor revisions	02.01.2019
Image3_1.2-PDMQ	ICM_AC	Further minor revisions	04.01.2019
Image3_2.0-PDMQ	RCSI_AA	Substantive changes	10.01.2019
Research Data report_1.0-PDMQ	RCSI_AA	Original document	01.01.2019
Research Data report_1.1-PDMQ	ICM_AC	Revisions and comments	02.01.2019
Research Data report_2.0-PDMQ	RCSI_AA	Substantive changes	10.01.2019

For proteomics datasets generated at RCSI, PRIDE (Proteomics IDEntifications) Archive convention (https://www.ebi.ac.uk/pride/archive/) will be implemented. This will facilitate the submission of published data via ProteomeXchange.

2.1.3 Metadata:

The extra information that surrounds research data, allows people to find, access and ultimately reuse data. Metadata standards provide specific data fields or elements to be used in describing data for a particular use. Some research fields have predefined metadata standards. Metadata will include clear labelling of data type, versions, and dates to ensure the data is understood and contextualized for future reuse. With respect to metadata standards, FAIRsharing.org is a useful database of different standards and how they relate to the final intended repositories and some publishers. For example, GEO repository has specific metadata requirements which are provided in a template prior to data upload (GEO repository metadata requirements). Being aware of such requirements from the initiation of a project ensures a smooth workflow within the final stages of data sharing. The Digital Curation Centre (DCC) recommend several standards for metadata schemata dependent on final intended repository upload recommended metadata standards.

Metadata associated with the data generated during the PD-MitoQUANT project will be embedded through the use of the PDMQ identifier, indicating the cell/tissue type, the nucleic acid extracted etc. Inclusion of the standard operating procedure (SOPs) followed for DNA and RNA extraction from samples and the SOP for sample processing during sequencing will be included. Metadata for 'omics datasets will be collected using the ISA Framework ISA (Investigation, Study, Assay) Framework. For derived data, the processing steps, software (code, parameters, version), OS environment, referenced dataset will be documented either on internal servers and/or in the PD-MitoQUANT Sharepoint data sharing platform. With respect to digitally captured images of immunohistochemistry (IHC), immune markers will be indicated on each captured image; and for IHC data held in raw TIFF files, images will also record the microscope settings at the time of capture. All relevant metadata surrounding protocols will be recorded in publications, including supplementary methods.

For *in vivo* experiments "standard" metadata should include mouse strain, model used, timing of injections, routes of injection, method of tissue collection, concentration of any injected agents (and the supplier/catalogue number, if commercially available) and animal monitoring (in all cases). The aforementioned (meta)data will be collected in text format and reported within, or as supplementary material to, any published data from the PD-MitoQUANT project. In subsequent updates of the DMSP, additional information on metadata will be added.

2.2 Making data openly accessible:

2.2.1 Accessibility:

The Consortium Agreement (CA) data retention guidelines will attempt to strike a balance between the scientific need to retain research data and the long-term cost of doing so. Subsequent to any decision made with regards to data sharing, recommendations from the EU-funded OpenAIRE project will be considered, such as depositing data in the EC commissioned cost-free repository.

The Data Steward will work with partner Data Coordinators and Processors, to determine the details of how PD-MitoQUANT data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling re-use, and definition of whether access will be widely open or restricted to specific groups. Long term curation of the data, and provision of access to authorised bona-fide stakeholders (researchers, clinicians etc.), is the mandated responsibility of the selected repository.

Notably, data will be stored at the respective collection sites and corresponding institutions. Further data storage and backup can be provided by the Data Coordinator (RCSI) IT support for this project. Data will be made available to all partners that contributed to the data generation and data analysis after completion of the program for an indefinite period of time. Only designated members of the consortium will have user and/or administrative privileges to view the data. In addition, knowledge derived from the research using the data may be brought forward to commercial purposes as appropriate, and this process will be regulated by the Grant Agreement and the Consortium Agreement, in accordance with any generally valid legislation and regulations.

Data extraction, processing and inputting will be undertaken by the in-country researchers, who

will also be responsible for routine supervision of the local dataset development including collecting and transcribing data, with the Data Coordinator (RCSI) supporting as necessary.

At the end of the project, the final (meta) data files from the PD-MitoQUANT project will be transferred and deposited in an appropriate data repository, following assessment by the PD-MitoQUANT STEER committee with respect to IP potential. These data repositories' security policies have been written according to best practices, for indefinite data retention and to ensure that the research community has long-term access to the data. In addition to the research community, we expect these data will be used by practitioners and policymakers. Data will be securely stored for a designated length of time in line with statutory and legal obligations imposed by the Data Coordinator (RCSI) before being deleted or disposed of in a secure manner in line with Data Protection guidelines.

2.2.2 Expected Data Repositories:

Modelling datasets will be deposited in GitHub or similar, whilst experimental data for modelling will be deposited in DRYAD or similar, following assessment for publication and IP potential in accordance with the Consortium Agreement. Mass Spectrometry (meta)data and other proteomics data will be deposited in the PRIDE PRoteomics IDEntifications (PRIDE) archives or similar. The PRIDE database is a centralized, standards compliant, public data repository for proteomics data, including protein and peptide identifications, post-translational modifications and supporting spectral evidence. PRIDE is a core member in the ProteomeXchange (PX) consortium, which provides a single point for submitting mass spectrometry based proteomics data to public-domain repositories. Datasets are submitted to PRIDE via ProteomeXchange and are handled by expert biocurators.

Preclinical RNASeq data will be deposited in NCBI's Gene Expression Omnibus (GEO) or similar publically available repositories.

Imaging data will be made available in the DRYAD repository when associated with scientific peer review publications generated as a result of work carried out during the PD-MitoQUANT project. DRYAD is an international disciplinary repository of data underlying scientific and medical publications. Dryad enables authors, journals, societies and publishers to facilitate data archiving at the time of publication, when the data are readily available. Data in Dryad receives a permanent, unique Digital object identifier (DOI), which can be included in the published article so that readers are able to access the data. Authors can archive data in Dryad and be assured of its preservation, while satisfying journals' and research funding agencies' mandates to disseminate their research outputs (www.journals.uchicago.edu/doi/10.1086/650340).

2.3 Making data interoperable:

For long term storage, interoperability and archiving of data generated during the PD-MitoQUANT Project it is important to address the format within which all generated data should be stored. All digital information is designed to be interpreted by computer programs to make it understandable; it is therefore inherently dependent on software. All digital data are thus at risk of becoming inaccessible should the hardware or software environments they depend on become obsolete. To guarantee long-term data access and usable data, data generated during the PD-MitoQUANT project should also be converted to **standard formats** such as those outlined in Table 3 that are then suitable for data interchange and transformation, adhering to FAIRification Data sharing. In brief, PD-MitoQUANT consortium members are requested to store all generated data as follows: **Image data as .TIFF, Data in spreadsheets will be stored as .csv, Data in free text documents will be stored as .txt.**

These formats are platform agnostic and should support future access and reuse. Any data which has to be stored in a proprietary format will have the necessary software (including version number) noted in an associated information text file. Additional acceptable open or standard formats include OpenDocument Format (ODF), ASCII, tab-delimited format, comma-separated values or XML; instead of proprietary ones (such as Microsoft Rich Text Format, Microsoft Excel and SPSS). Although proprietary formats are widely used and likely to be accessible for a reasonable time, this time frame is not unlimited and therefore an additional, or alternative, format is also required. In this regard, while PD-MitoQUANT consortium researchers use the most

suitable data formats and software according to planned analyses, once data analysis is completed and data are prepared for storing, researchers should save an additional version of their data by converting their research data to standard, interchangeable and longer-lasting formats to avoid being unable to use the data in the future. Similarly for backups of data, standard formats need to be saved. Implementing such protocols routinely from the initiation of the project eases conversion processes at later stages. Additional forms of acceptable open and standard formats are listed in Table 3.

2.4 Increase data re-use (through clarifying licenses)

2.4.1 Third Party Access:

All data published from the PD-MitoQUANT project will be publically available and therefore will allow third party access. Prior to publication, a potential user who wants access to the data will be instructed to submit a request to the Data Coordinator including the applicant's full name, associated institution/organization, departmental email address, and a detailed description of intended use. All requests will be reviewed by the Project Coordinator Data Steward (RCSI), Project Coordinator Data Protection Officer (DPO) and the PD-MitoQUANT Steering Committee on a case-by-case basis, and will be subject to IP potential evaluation prior to data release

If approved, the requester will receive an email with a unique link to verify the email address. Once the email address has been verified, the applicant will be asked to agree and sign the terms and conditions of access (e.g. the researcher shall use the dataset only for non-commercial research and educational purposes). The Data Coordinator (RCSI) will make no representations or warranties regarding the dataset, including but not limited to warranties of non-infringement or fitness for a particular purpose. The applicant researcher accepts full responsibility for his or her use of the dataset and shall indemnify the Data Coordinator, including their employees, Trustees, officers and agents, against any and all claims arising from applicant researcher's use of the dataset, including but not limited to applicant researcher's use of any copies of copyrighted dataset that he or she may create from the dataset. The applicant researcher may provide research associates and colleagues with access to the Dataset provided that they first provide written agreement to be bound by these terms and conditions.

The Data Coordinator (RCSI) reserves the right to terminate the researcher's access to the database at any time. If the applicant researcher is employed by a for-profit, commercial entity, the researcher's employer shall also be bound by these terms and conditions, and the researcher hereby represents that he or she is fully authorized to enter into this agreement on behalf of such employer. The law of the Data Coordinator's country (Ireland) shall apply to all disputes under this agreement. Records will be maintained to keep track of entities using the consortium dataset. These third parties will not have rights for redistribution or publishing of the dataset, partially or in full, requiring the reference to the dataset in any publication or derived product.

Table 3: Guidance on file formats (www.ukdataservice.ac.uk)

Type of data	Recommended formats	Acceptable formats
Tabular data with extensive metadata variable labels, code labels, and defined missing values	SPSS portable format (.por) Delimited text and command ('setup') file (SPSS, Stata, SAS, etc.); structured text or mark-up file of metadata information, e.g. DDI XML file	proprietary formats of statistical packages: SPSS (.sav), Stata (.dta), MS Access (.mdb/.accdb)
Tabular data with minimal metadata column headings, variable names	comma-separated values (.csv) tab-delimited file (.tab) delimited text with SQL data definition statements	delimited text (.txt) with characters not present in data used as delimiters; widely-used formats: MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb), dBase (.dbf), OpenDocument Spreadsheet (.ods)
Geospatial data vector and raster data	ESRI Shapefile (.shp, .shx, .dbf, .prj, .sbx, .sbn optional) geo-referenced TIFF (.TIFF, .tfw) CAD data (.dwg) tabular GIS attribute data Geography Markup Language (.gml)	ESRI Geodatabase format (.mdb); MapInfo Interchange Format (.mif) for vector data; Keyhole Mark-up Language (.kml); Adobe Illustrator (.ai), CAD data (.dxf or .svg); binary formats of GIS and CAD packages
Textual data	Rich Text Format (.rtf) plain text, ASCII (.txt) eXtensible Mark-up Language (.xml) text according to an appropriate Document Type Definition (DTD) or schema	Hypertext Mark-up Language (.html) widely-used formats: MS Word (.doc/.docx) some software-specific formats: NUD*IST, NVivo and ATLAS.ti
Image data	TIFF 6.0 uncompressed (.tif)	JPEG (.jpeg, .jpg, .jp2) if original created in this format; GIF (.gif); TIFF other versions (.tif, .tiff); RAW image format (.raw); Photoshop files (.psd); BMP (.bmp); PNG (.png); Adobe Portable Document Format (PDF/A, PDF) (.pdf)
Audio data	Free Lossless Audio Codec (FLAC) (.flac)	MPEG-1 Audio Layer 3 (.mp3) if original created in this format Audio Interchange File Format (.aif) Waveform Audio Format (.wav)
Video data	MPEG-4 (.mp4) OGG video (.ogv, .ogg) motion JPEG 2000 (.mj2)	AVCHD video (.avchd)
Documentation and scripts	Rich Text Format (.rtf) PDF/UA, PDF/A or PDF (.pdf) XHTML or HTML (.xhtml, .htm) OpenDocument Text (.odt)	plain text (.txt) widely-used formats: MS Word (.doc/.docx), MS Excel (.xls/.xlsx) XML marked-up text (.xml) according to an appropriate DTD or schema, e.g. XHMTL 1.0

2.4.2 Copyright and Intellectual Property:

Aspects related to copyright and Intellectual Property Rights (IPR) issues have been established in the Consortium Agreement between each beneficiary. The copyright and intellectual property for all works derived from the primary data resides primarily with the laboratory where the data was collected and with the collaboration that has initially collected the data set (unless otherwise specified in the Consortium Agreement). Any subsequent use of such data (re-analysis, analysis of side channels, etc.) needs to be done in close collaboration with the original collaborators and agreement for such use must be sought before any use of the data is made.

3. Allocation of resources

In order to ensure that the data generated is FAIR and compliant, researcher time has been allocated in the programme budget to cover the costs of preparing data and documentation for archiving.

3.1 Data Management Roles:

The Data Coordinators for the PD-MitoQUANT project are RCSI (Coordinator), including the PD-MitoQUANT Research Project and Data Manager, and the FAIR Data Steward (Dr. Stephen Madden, RCSI). RCSI's established Institutional Data Protection Officer (DPO) (Mr. Dónall King) will provide opinion or confirmation regarding the compliance of all PD-MitoQUANT data collection and processing with EU and national data protection legislation. This will be retained on file and submitted to the Commission upon request. Each partner who generates data in any format is referred to as a Data Processor. The Data Processor in some instances may also be the researchers themselves, depending on the partner institutional framework. Each PD-MitoQUANT consortium partner Data Processor will provide the Data Coordinator (RCSI) with the contact name and details of their institutional/company Data Manager and/or Data Protection Officer (DPO).

3.2 Resources for long term preservation:

For data that will be made available to all partners such as shared protocols, analyzed RPPA data and systems-based network analyses, data will ultimately be housed at RCSI. Data will be preserved locally and backed up through the RCSI IT Department every four weeks. The ownership of uploaded data and models will remain with the groups that produced them. Standard Operating Procedures (SOPs) for analytical procedures will also be stored and linked to the data. Sharing of data generated during the project for the scientific community will use the following community data and tool repositories that ensure long term preservation of data:

- Models and model codes: GitHub or similar
- Experimental data: DRYAD
- Proteomics data: PRIDE or similar
- Raw data sets agreed to be made publicly available will also be submitted to the appropriate databases (e.g.: NCBI Gene Expression Omnibus, NCBI Sequence Read Archive).

4. Data security

Guidelines for data security and personal data protection will be followed, which is according to the recommendation set forth in the Code of Practice for Reuse of Medical Data in Scientific Research Projects developed in the eTRIKS, currently under review by the European Data Protection Supervisor for endorsement throughout Horizon 2020 Health Research related activities. To protect data storage processing, the following safety measures will be undertaken:

- Compliance with the General Data Protection Regulation (GDPR) (EU) 2016/679 legislation from 25th May 2018.

- Regular security monitoring and reporting any irregularities to the Data Steward (RSCI) and Data Coordinator (RCSI)
- Reporting of data security incidents including personal data breaches to the Data Steward (RSCI), and also to the Data Protection Officer (DPO) of the related dataset within two working days after becoming aware of it.

The main risk to data security are potential violations of the SOP defined for data sharing, e.g.:

- Sending sensitive data via email instead of uploading to servers using encrypted communications (SCP, HTTPS).
- Disclosing or sharing credentials to non-authorized parties.
- Users left their job/role during the project, but this information is not communicated in time to the server-admin team.

The PD-MitoQUANT consortium understands its mandatory obligation under the GDPR to report data breaches to the Office of the Data Protection Commissioner within 72 hours where there is a risk to the rights of the data subject(s). The PD-MitoQUANT consortium understands the implementable fines and measures for noncompliance and the right to compensation of the data subjects.

4.1 Storage, protection and destruction:

Upon finalization of the RCSI Data storage platform, all research data generated during the course of the PD-MitoQUANT Project may be stored on a server located at RCSI, within Ireland; protected by strong encryption and in compliance with GDPR and data protection law. The data files will be managed, processed, and stored in a secure environment including lockable computer systems with passwords, firewall system in place, power failure or surge protection, virus/malicious intruder protection and by controlling access to the files with encryption and password protection. Data backup on RCSI servers is undertaken every 4 weeks and technical expertise is available from the RCSI IT Services for the setup of the files within dedicated password protected folders. In addition, support and training are available should this be required. Data that is archived will be migrated to Azure with data being processed and stored on the Isilon.

Secure storage and protection of research data generated by PD-MitoQUANT consortium partners remains the responsibility of the individual partners. Each partner should ensure that their processes for data storage, encryption and protection are compliant with GDPR regulations and adhere to best practice guidelines. In addition to local backup and archiving requirements, data from partners which is uploaded to the PDMQ-data sharing platform will be stored and backed up accordingly.

Research data must be retained and disposed of securely according to the relevant retention and disposal schedule, in accordance with legal, ethical, research funder and collaborator requirements and with particular concern for the confidentiality and security of the data. Research data that underpins published results or is considered to have long-term value should be retained, subject to informed consent to do so, where relevant. In the absence of the other provisions, the default period for research data retention is 10 years from date of last requested access.

PD-MitoQUANT will implement research integrity and data management practices that apply appropriate protections and which recognize the legal, ethical and commercial constraints that may impinge upon the release of research data.

5. Ethical aspects

The processes and documentation for obtaining informed consent for the use of human cells and tissues during the PD-MitoQUANT project will be comprehensive. All data collected and generated will be in compliance with good practice principles for undertaking research including international/EU conventions and declarations, including the Declaration of Helsinki (Oct 2003 and in its latest versions), the Council of Europe Convention of Protection of Human Rights and Biomedicine, Clinical Trials Directive 2001/20/EC and Directive 2001/83/EC, Good Clinical Practice Directive (2005/28/EC) and others. Activities involving the processing of personal data for the programme will comply with applicable national Data Protection Laws which implement the

General Data Protection Regulation (GDPR) (EU) 2016/679.

For clinical data collected with respect to human iPSCs and Parkinson's disease patient tissue samples, partner organizations (CNR and PUK, respectively) will ensure that all patient samples will be processed in line with the informed consent collected at individual institutes. Ethical approval, patient consent and patient information leaflets for these studies will be held on file by the Data Coordinator (RCSI) and submitted to the funding agency upon request (WP4 D4.1). In all cases where human clinical data is used, the key for decoding will stay with the participating institution-hospital/biobank and will not be made available outside the hospital/institution, i.e. this information will not be shared with the consortium members. This coding includes not only the removal of personal identifiers and information, but also of clinical information that could potentially be used to identify a single individual or to expose transmissible disease susceptibilities. Patients will not be identified in any reports on this study.

The Data Coordinator (RCSI) will be responsible for obtaining copies of the formal consent for data preservation, sharing and reuse and will make these documents available on Sharepoint for consortium members only, in order to facilitate local ethics applications to receive and undertake research on these human-derived samples. Also, the Data Coordinator will ensure that a Data Protection Impact Assessment (DPIA) has been carried out at applicable institutions involved in the PD-MitoQUANT consortium, whereby human patient clinical information is protected and has been collected and pseudonymised/anonymized, as required by GDPR legislation. The retention of a DPIA document will help ensure patient data is stored, transferred and handled securely in accordance with GDPR. The Data Coordinator (RCSI) will monitor data production, sharing and storage to ensure compliance with the relevant DPIA and that all guidelines and recommendations within the DMSP have been translated to project development.

6. Other

All PD-MitoQUANT partners will comply with their own in-house and/or departmental procedures for data management regarding back-up, storage etc. Ethical standards and guidelines of Horizon 2020 will be rigorously applied. The use and transfer of PD-MitoQUANT personal data generated as a result of RNA sequencing is covered by the EU Data Protection Directive 95/46/EC, which ensures the protection of individuals, and EU Regulation 2016/679 of the European Parliament and of the Council of 27 April 2016 which applies in all member states from 25 May 2018, will be taken into account in the work of the project in order to ensure continuous compliance.

6.1 GDPR regulations re use of human tissues or cells:

The datasets collected in this project will initially only be used for the research topic defined in the research proposal and within the scope of the informed consent. Research undertaken within PD-MitoQUANT falls into the category of health research and the definitions within this category that are specifically related to PD-MitoQUANT are:

- Research with the goal of understanding the normal and abnormal functioning, at the molecular, cellular, organ system and whole-body level.
- Research that is specifically concerned with innovative strategies, devices, products or services for the diagnostic, treatment of prevention of human disease or injury.
- Research with the goal of improving the diagnosis and treatment of human disease and injury and of improving the health and quality of life of individuals.
- It includes a necessary action to establish whether an individual may be suitable for inclusion in the research where the processing of any personal data for that purpose shall be subject to the Act and the Dara Protection regulation.

Regarding the type of data that will be used, sensitive personal data from human clinical samples will be pseudonymised before such data will be shared with the consortium. **It is important to note that the patient will not be identifiable within the PD-MitoQUANT consortium as keys to patient identity remain at the sample collection site or biobank.** Data arising from iPSC experimental research will be considered non-personal data and will be managed by PD-MitoQUANT as such.

RNA sequencing data obtained from human iPSCs or human PD tissues will be considered

personal data. **The lawful basis for processing such personal data in PD-MitoQUANT falls within Article 6 and Article 9 of GDPR which provide the legal basis for processing personal data and sensitive personal data respectively.** The specific legal bases for processing data and sensitive data within PD-MitoQUANT are as follows:

- Article 6.1(a): The data subject has given consent to the processing of his or her personal data for one or more specific purposes;
- Article 6.1(f): Processing is necessary for "Legitimate interests", which includes research;
- Article 9.2(a): the data subject has given explicit consent to the processing of those personal data for one or more specified purposes, except where Union or Member State law provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject;
- Article 9.2(j): Processing is necessary for "Scientific research purposes" (in which case **data must be pseudonymised**).

6.2 PDMQ Data Management Policies:

Good research data management will enable PD-MitoQUANT researchers to meet funder, ethical, legal standards and responsibilities. It also ensures that research data is accurate, complete, authentic and reliable, stored securely, preserved where necessary and accessible as required. It also enables access by other researchers who could use the data, thus maximizing the effectiveness of PD-MitoQUANT research funding, raising awareness of PD-MitoQUANT research and researchers and demonstrating alignment with Open Science, Open Access and the FAIR data principles.

This policy applies to all consortium members engaged in research, including staff and research students, and those who are conducting research on behalf of PD-MitoQUANT. It applies to all research irrespective of funding. PDMQ recognizes research data as a valuable institutional asset, and the role of research data management in underpinning research excellence and integrity. Researchers have the primary responsibility to ensure research data will be managed in line with funder requirements as well as relevant regulations and legislation.

PD-MitoQUANT commits to: (a) Disseminating information amongst its academics about the requirements under this policy framework and under policies of the PDMQ funders in relation to research data; and (b) Developing infrastructure and training to promote best practice in data management amongst its academics, to acknowledge its obligations and achieve compliance with this policy framework and with its funders' data policies.

Research data must be: (a) As compatible as possible with the FAIR data principles, as open as possible and restricted as necessary; (b) Secure and safe with appropriate measures taken in handling sensitive, classified and confidential data; (c) Stored in a manner that is compliant with legal obligations, institutional policy and, where applicable, the requirements of funding bodies; and (d) Preserved for its life-cycle with the appropriate high-quality metadata.

6.3 Information Security Policy:

The Information Security policy, available from the Data Coordinator (RCSI), provides a framework in which security threats to Information Systems can be identified and managed on a risk basis and establishes terms of reference, which are to ensure uniform implementation of information security controls throughout the consortium. The objectives of this policy are to:

- Ensure that information is created and maintained in a secure environment;
- Ensure that all of the consortium partner's computing facilities, programs, data, network and equipment are adequately protected against loss, misuse or abuse;
- Create awareness that appropriate security measures must be implemented as part of the effective operation and support of Information Security;
- Ensure that all users understand their own responsibilities for protecting the confidentiality and integrity of the data they handle;
- Ensure all consortium partner owned assets have an identified owner /administrator;
- Ensure that all users are aware of and fully comply with the relevant Irish and European Community legislation.

7. Proposed update schedule for Data Management and Sharing Plan

Proposed Schedule for Future Versions of DMSP				
Version	**Publication Date**	**Submission Date**		**Work package and deliverable number**
Initial	1st July 2019	1st August 2019	6 months	WP3 D3.3
Detailed	1st July 2020	1st August 2020	18 months	WP3 D3.4
Final	1st January 2022	1st February 2022	36 months	WP3 D3.8

Conclusion

This deliverable has provided confirmation that all relevant Data Management and Sharing documents required for generating, using, sharing or storing (meta) data during the course of PD-MitoQUANT were established and agreed upon by all consortium members prior to the start of the associated research activities. Relevant documentation referred to within the Data Management and Sharing plan are kept on file and will be provided to the IMI JU on request.