DATA MANAGEMENT PLAN

Grant Agreement number	22HLT07
Project short name	NEuroBioStand
Project full title	Standardisation of measurements of neurodegenerative disease biomarkers
Data management plan	1st ☒ 2nd ☐ 3rd ☐

TABLE OF CONTENTS

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EURAMET. Neither the European Union nor the granting authority can be held responsible for them.

The project has received funding from the European Partnership on Metrology, co-financed from the European Union's Horizon Europe Research and Innovation Programme and by the Participating States.

European Partnership Co-funded by the European Union

METROLOGY PARTNERSHIP EURAMET

1 Data management plan

1.1 Data summary

Questions	Answers
1 Will you re-use any existing data and what will you re-use them for? State the reasons if re-use of any existing data has been considered but discarded.	This project will re-use: • Internal data from the participants • Publicly available data These data will be used for the following purpose: • Validation of the project's results
2 What types and formats of data will the project generate or re-use?	• numerical data in CSV format • graphics/images in jpeg, tiff or png • tables in opj, xlsx or csv • text files (e.g. protocols and procedure description) in docx, pdf or txt • raw data from different LC-MS software
3 What is the purpose of the data generation or re-use and its relation to the objectives of the project?	<u>Purpose of the data generation or re-use</u> The data generated and re-used will be from measurements, calibrations, comparisons and validations. They will be used in meeting the project's objectives and in conference and peer-reviewed publications. <u>Data generated in relation to the objectives of the project</u> Data will be generated by the consortium in order to meet objectives 1 - 4. Measurement and calibration data will result from objectives 1 to 4 and comparison and validation data from objectives 1, 2, 3 and 4. Data from questionnaires and market surveys will be used to support end-user uptake (objective 5). <u>Data re-used in relation to the objectives of the project</u> Measurement, calibration, comparison and validation data will be re-used by the consortium in order to meet objectives 1 to 5.
4 What is the expected size of the data that you intend to generate or re-use?	The overall size of the data is expected to be in the range 1 to 2 TB. This number will be updated in the second version of the DMP.
5 What is the origin/provenance of the data, either generated or re-used?	<u>Data generated in the project</u> The data generated will be from measurements, calibrations, comparisons and validations. <u>Re-used data</u> The existing data will originate from several sources, which will include:participant's pre-existing data, data from the scientific literature, real-world measurement data and data from simulation experiments.
6 To whom might your data be useful ('data utility'), outside your project?	The data will be suitable for use by other research groups working on the following topics: neurodegenerative diseases, biomarkers, protein analysis, mass spectrometry. It will also be useful for standards committees including IFCC-WG-BND and IFCC-WG-CMT. The data might be useful to: • Stakeholders from industry: IVD providers, biopharma, mass spectrometry manufacturers • Standardisation bodies: IFCC-WG-BND, IFCC-WG-CMT, CCQM PAWG, JCTLM

	• EQA providers • NMIs/DIs • Other scientists working in the field: clinicians in neurodegenerative diseases • Economic actors: • Decision makers: … • Citizens

1.2 Findable, Accessible, Interoperable and Re-usable (FAIR) Data

1.2.1 Making data findable, including provisions for metadata

Questions	Answers
7 Will data be identified by a persistent identifier?	All data, which will be uploaded into a trusted repository during and after the project termination will receive a DOI as a persistent identifier. Each partner, who will upload data, is responsible for obtaining an appropriate DOI. Furthermore, they will be made available in a future version of the project DMP.
8 Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.	The metadata created for all of the project's deposited datasets will be open under a Creative Commons Public Domain Dedication (CC 0) or equivalent (to the extent legitimate interests or constraints are safeguarded), in line with the FAIR principles (in particular machine-actionable) and provide information at least about the following: datasets (description, date of deposit, author(s), venue and embargo); the European Partnership on Metrology funding; grant project name, acronym and number; licensing terms; persistent identifiers for the dataset, the authors involved in the action, and, if possible, for their organisations and the grant. Where applicable, the metadata will include persistent identifiers for related publications and other research outputs.
9 Will search keywords be provided in the metadata to optimise the possibility for discovery and then potential re-use?	Yes, the following search keywords will be provided in the metadata to optimise the discovery and potential re-use of the deposited datasets: neurodegenerative diseases, tau, NfL, GFAP, standardisation, biomarkers, traceability. Further keywords will be identified in future DMP versions.
10 Will metadata be offered in such a way that it can be harvested and indexed?	Zenodo complies with FAIR principles (https://about.zenodo.org/principles/). The metadata are indexed in a searchable resource. Metadata are licensed under CC0, except for email addresses. It is planned, if it is technically feasible, to make all metadata exportable via a specific protocol and harvested.

1.2.2 Making data accessible

Questions	Answers
Repository:	Zenodo
11 Will the data be deposited in a trusted repository?	The data and associated metadata, documentation and code will be deposited in the trusted open access repository Zenodo (https://zenodo.org). Further repositories can be proposed by any participant and will be reviewed and checked for compliance with the EU

Questions	Answers
	requirements. Only if they have passed this assessment, then they will be made available for uploading any data.
12 Have you explored appropriate arrangements with the identified repository where your data will be deposited?	No, the data will be uploaded via a standard procedure and require no special arrangements.
13 Does the repository ensure that the data are assigned an identifier? Will the repository resolve the identifier to a digital object?	Yes, Zenodo will assign an identifier (DOI) to each of the project's deposited datasets. The repository will resolve the identifier to a digital object.
Data:	
14 Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.	All of the data that are needed to validate the results presented in scientific publications will be made openly available as the default unless there is a specific reason not to publish the data. <u>Datasets which cannot be shared – voluntary restrictions</u> Other data may be made available on a case-by-case basis if it is relevant for third parties. The following data will not be made publicly available: • Data obtained with the permission of third parties, but the third parties have not agreed to make the data publicly available. • Data that discloses the identity of a manufacturer. • Data that compromises the protection of a participant(s) intellectual property. The level of data made available will also be considered, for example, pre-processed data will not be provided unless there is a clear reason for doing so. <u>Datasets which cannot be shared - legal / contractual reasons</u> All of the data from the project will be made available, with the exception of market or customer survey data, which are commercially sensitive and cannot be shared.
15 If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.	The data used in scientific publications, posters and oral communications will be made available for re-use as soon as is reasonably possible. Some of the data are expected to be subject to an embargo period of 18 months whilst a patent application is pending.
16 Will the data be accessible through a free and standardised access protocol?	Yes, Zenodo provides well described conditions for free and standardised access (see http://about.zenodo.org/policies/).
17 If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?	There are no restrictions on the use of the published data, but users will be required to acknowledge the project and the source of the data in any resulting publications, according to the latest version of the CC-BY license.

Questions	Answers
18 How will the identity of the person accessing the data be ascertained?	There is no need to ascertain the identity of persons accessing the data.
19 Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?	This consortium will not establish a Data Access Committee. The coordinator, with support from the participants, will have overall responsibility for the management of data/research outputs and quality assurance.
Metadata:	
20 Will metadata be made openly available and licensed under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?	In Zenodo, metadata are licensed under CC0, except for email addresses. All metadata can be exported via a specific protocol and harvested.
21 How long will the data remain available and findable? Will metadata be guaranteed to remain available after data are no longer available?	The data will remain available and findable for the lifetime of the Zenodo repository, which is expected to be a minimum of 20 years. If data are withdrawn from Zenodo, the DOI and the URL of the original object are retained. In case of closure of the Zenodo repository, best efforts will be made by Zenodo to integrate all content into suitable alternative institutional and/or subject based repositories
22 Will documentation or reference about any software be needed to access or read the data and will this be included? Will it be possible to include the relevant software (e.g. in open source code)?	Some of the data are in a common format and can be read using widely available software (open source or commercial). Mass spectrometry data can be read using specialised scientific software (open source or commercial).

1.2.3 Making data interoperable

Questions	Answers
23 What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?	The datasets will use the trusted repository's basic metadata schema for administrative data, which is compliant with the recommended standards used by DataCite (https://search.datacite.org/) and OpenAIRE (https://www.basesearch.net/). For individual datasets, the following discipline-specific vocabularies, standards, formats, and methodologies will be used: 1. GUM (procedure; subject-independent). 2. ISO 9001 (QM procedure; subject-independent) 3. ISO/IEC Guide 99:2012 (VIM 3) 4. ISO/IEC 17034 and 17025
24 In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly	Mapping will not be required as the terminology used will be chosen to be compatible with the existing literature.

used ontologies? Will you openly publish the generated ontologies or vocabularies to allow their re-use, refinement or extension?	
25 Will your data include qualified references[1] to other data (e.g. other data from your project, or datasets from previous research)?	Yes, the project's datasets that will be deposited in the chosen repository (e.g. Zenodo) will include qualified references to other datasets from the same project and from previous research.

1.2.4 Increase data re-use

Questions	Answers
26 How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?	If needed, a short README file will be provided together with the data, in order to enable data analysis and to facilitate data re-use.
27 Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard re-use licenses, in line with the obligations set out in the Grant Agreement?	The data will either be licensed under the latest available version of the Creative Commons Attribution International Public License (CC BY) or a license with equivalent rights as set out in the Grant Agreement. Users will be required to acknowledge the consortium and the source of the data in any resulting publications. Alternatively, the Creative Commons Public Domain Dedication License (CC 0) will be used.
28 Will the data produced in the project be useable by third parties, in particular after the end of the project?	Any data published in open-access journals will be usable by third parties after the datasets have been deposited in Zenodo. The data that do not relate to peer-reviewed publications will be made available for re-use on a case-by-case basis.
29 Will the provenance of the data be thoroughly documented using the appropriate standards?	Yes, the provenance and context of the data will be thoroughly documented to meet relevant standards using the the Dublin Core Standard. Data will be accompanied by information on how they were captured, processed, analysed, and validated. Other information that enables interpretation and use will also be provided.
30 Describe all relevant data quality assurance processes.	Data quality will be assured through several quality assurance procedures: • Repeated and comparison measurements. • Adherence to standards for data recording. • Use of controlled vocabularies and standard terminology. • Metrological characterisation of the measurement set-ups. • Validation of the data collected. • Provision of test results along with the data. • Peer-review of publications based on the data.

[1] A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/)

Questions	Answers
31 Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.	Allocation of resources The estimated costs for making the (data and) other research outputs FAIR are 1000 € (personnel costs) (see question 34). The costs for making other research outputs FAIR are included in the project's budget and will be claimed if compliant with the Grant Agreement's conditions. The coordinator will also have overall responsibility for managing other research outputs (see question 36). Where feasible, long-term preservation will be ensured by depositing the other research outputs in repositories. The coordinator will decide on a case-by-case basis on which other research outputs will be deposited and for how long. Security of other research outputs All participants are either accredited to, or work in compliance with, the ISO 17025 standard on the "General requirements for the competence of testing and calibration laboratories". The participants will store other research outputs on their organisations' networks, which are protected by firewall, backups etc. Other research outputs will also be stored in the project's SharePoint environment, with password-protected login. Deposition in public repositories will provide additional security as they have multiple replicas in a distributed file system which is backed up on a nightly basis. This project will not generate sensitive other research outputs. The other research outputs will be safely stored in open access repositories. Ethical aspects There are issues that could impact on the sharing of other research outputs. • Information relating to other research outputs acquired from third parties, e.g. manufacturers, will not be shared without their explicit consent. • Information relating to other research outputs collected by the consortium at commercial sites will not be shared without the site owner's explicit consent. Ethical issues will be addressed as the project will prepare an ethics report. The project will not share other research outputs with identifiable personal information. Sensitive information relating to the other research outputs will be collected, separated as soon as possible and kept secure.

1.3 Other research outputs

Questions	Answers
32 In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models,	The new calibration methods, and protocols produced by the project will be stored in the Protocol Exchange repository. The management of the IP issues surrounding the new materials that will be developed in the project have been planned in the project's consortium agreement. The consortium may seek patent protection on new candidate reference materials. This will be stated in the next versions of the DMP. This project will only re-use existing data and will not re-use any other research outputs.

Questions	Answers
etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).	
33 Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.	As far as possible, the FAIR data approaches specified in questions 7-30 above will be applied to the management of this project's other research outputs. This commitment will be met by placing the new calibration methods, and protocols, in a trusted repository and by eventually patenting the new materials that will be developed in the project in line with the requirements of the project's consortium agreement.

1.4 Allocation of resources

Questions	Answers
34 What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ?	The estimated costs for making the data and other research outputs Findable, Accessible, Interoperable and Re-usable (FAIR) are 100 € (personnel costs). These costs have been kept to a minimum by using a free repository (Zenodo) and by making only relevant data and other outputs FAIR.
35 How will these be covered? Note that costs related to research data/output management are eligible as part of the European partnership on metrology grant (if compliant with the Grant Agreement conditions).	The costs for making the data FAIR are included in the project's budget and will be claimed if compliant with the Grant Agreement's conditions.
36 Who will be responsible for data management in your project?	The coordinator will have overall responsibility for data management and will be responsible for coordinating updates to the data management plan. The coordinator will be responsible for organising data backup and storage, data archiving and for depositing the data within the repositories (Zenodo).
37 How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?	Long term preservation will be ensured by depositing the data within repositories (Zenodo). There are no costs associated with the long-term preservation of the data in these repositories. The data will increase in value over time because of its fundamental impact in a wide range of applications. It will enable the technologies developed in the project to be taken up by the measurement supply chain and by standards including IFCC WG BND. These standards bodies will need access to the data to justify the robustness of future standards. The data will also be of value as it underpins the results of published datasets. The coordinator will decide on a case by case basis on what data will be kept and for how long.

1.5 Data security

Questions	Answers
38 What provisions are or will be in place for data security	Data recovery and secure storage

(including data recovery as well as secure storage/archiving and transfer of sensitive data)?	All participants are either accredited to, or work in compliance with, the ISO 17025 standard on the "General requirements for the competence of testing and calibration laboratories". The participants will store data on their organisations' networks, which are protected by firewall, backups etc. Data will also be stored in the project's SharePoint environment, with password protected login. Deposition in the Zenodo public repository will provide additional security as it has multiple replicas in a distributed file system which is backed up on a nightly basis. Transfer of sensitive data • This project will not generate sensitive data.
39 Will the data be safely stored in trusted repositories for long term preservation and curation?	Yes, the data will be safely stored in the Zenodo open access repository. Deposition in the Zenodo public repository will provide additional security as it has multiple replicas in a distributed file system which is backed up on a nightly basis. Zenodo is working towards ISO 16363 certification.

1.6 Ethics

Questions	Answers
40 Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics report(s) and the ethics section in the Annex 1.	There are issues that could impact on data sharing. • Data acquired from third parties, e.g. manufacturers, will not be shared without their explicit consent. • Data collected by the consortium at commercial sites will not be shared without the site owner's explicit consent. • The data from the market surveys will be made anonymous to comply with the General Data Protection Regulation (GDPR). Ethical issues will be addressed as the project will prepare and submit a report on the Dual Use of the project's results.
41 Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?	Informed consent for data sharing and long-term preservation will be included in the market and customer surveys, but the project has no plans to share data with identifiable personal information. If any sensitive data are collected they will be separated as soon as possible and kept secure.

1.7 Other issues

Questions	Answers
42 Do you, or will you, make use of other national / funder / sectorial / departmental procedures for data management? If yes, which ones (please list and briefly describe them)?	Data management will be compliant with: • The research data policy of the European Partnership on Metrology; • European laws about data security and the protection of privacy (e.g. GDPR); • Institutional guidelines; • Scientific community guidelines.

2 Open science: research data management

Statement	Put an X in the	Or, list any exceptions to this

	box to confirm	
All participants have adhered to the requirements of the project's GA and CA with respect to open science: research data management (GA Article 17 and its Annex 5) for this reporting period	☒	