FINAL PUBLISHABLE REPORT

Grant Agreement number	15HLT04
Project short name	NeuroMET
Project full title	Innovative measurements for improved diagnosis and management of neurodegenerative diseases

Project start date and duration:		July 2016, 36 months
Coordinator: Milena Quaglia, LGC	Tel: +44 20 8943 7589	E-mail: Milena.Quaglia@lgcgroup.com
Project website address: https://www.lgcgroup.com/our-programmes/empir-neuromet/		
Internal Funded Partners: 1 LGC, UK 2 INRIM, Italy 3 LNE, France 4 PTB, Germany 5 RISE, Sweden	External Funded Partners: 6 Charité, Germany 7 CHU Mpt, France 8 UCL, UK 9 UEA, UK	Unfunded Partners:
RMG: -		

TABLE OF CONTENTS

1 Overview

Neurodegeneration is an incurable, debilitating process which presents a growing global challenge due to the increasingly ageing population. Alzheimer's disease (AD) and Parkinson's disease (PD) are the two most common neurodegenerative diseases (NDD). Both involve the build-up of specific proteins in the brain and subsequent neurodegeneration leading to physical and mental impairment including dementia. Currently, there are no clinically validated, minimally invasive diagnostic tools which allow the early diagnosis and/or monitoring of disease progression in AD and PD patients and available therapeutics only offer transient symptomatic relief. This project aimed to develop reference measurement systems to support the major NDD patient needs including less invasive, more accurate diagnostic measurement, and improved treatment and anxiety monitoring. Measurement comparability, underpinned by SI traceability and uncertainty analysis, is an unmet requirement for regulatory approval of NDD biomarkers, patient centred outcome measures (PCOM), clinical thresholds and new therapeutic drugs. Therefore, the reference methods developed within the NeuroMET project, to support the production of calibrators and improve measurement comparability of established biomarkers in the NDD area will contribute significantly to advancement in the field. Furthermore, the uptake of the developed PCOM will have a significant positive impact on both NDD patient and clinical trial assessments.

2 Need

The "Implementation report on the Commission Communication on a European initiative on Alzheimer's disease and other dementias" (2014) highlighted the importance of early diagnosis in NDD. Established biomarkers of AD from cerebrospinal fluid (CSF) have been used to differentiate between subjects with mild cognitive impairment (MCI) who have progressed to AD and those MCI patients who have not. However, the lumbar puncture procedure for CSF sample collection is time consuming, invasive and therefore limited in terms of the possibility of widespread application for early AD diagnosis. Furthermore, although AD biomarker detection from less invasive diagnostic procedures such as blood samples and neuroimaging (i.e. Magnetic Resonance Imaging (MRI)) has significant advantages, the limits in measurement sensitivity and high measurement variability of recognised and novel biomarkers for AD and PD have constrained the development of clinical thresholds for NDD early diagnosis. Major impacts of NDD, particularly the decline in cognitive function, as well as increase in psychological symptoms (agitation, anxiety, etc.), can be captured in PCOMs. These PCOMs need to be correlated with the various biomarkers and objectively monitored and managed. The NeuroMET project addressed these needs for measurement improvement and standardisation of NDD non- and minimally invasive biomarker measurements to develop PCOMs, which were then validated by laboratory data, as well as by establishing reference measurement procedures to underpin measurement comparability. These are fundamental requirements both for the reliable development of minimally invasive early diagnostic and patient management procedures and also to support therapeutic discovery.

3 Objectives

This project combined the expertise of metrology laboratories together with clinicians and academics, in order to (i) overcome measurement issues currently constraining clinical innovation and uptake in NDD diagnosis and treatment, and to (ii) provide routes to directly translate research outputs into the clinic. Unusually for a metrology project, patients played a central role in the NeuroMET project, thus ensuring that metrology was correlated to patients' clinical status and is relevant to patient and clinical measurement requirements. The specific objectives of the project were:

1. **To establish patient cohorts** representative of AD (including MCI due to AD and dementia due to AD), PD neurodegeneration and a matched group of healthy control subjects.
2. **To develop and validate non-invasive magnetic resonance imaging (MRI) approaches** for *in vivo* characterisation of AD and MCI patients and healthy matched controls.
3. **To develop minimally invasive methods for early diagnosis and drug therapeutic monitoring**. This will be achieved by developing and applying conventional and innovative methods for the quantification of recognised and emerging biomarkers present in of the AD and PD patient cohorts.
4. **To improve NDD biomarker measurement comparability through the establishment of traceability chains**. This will be achieved through the development of reference methods and reference materials for

measurements, in CSF, of the established AD marker, tau protein (tau), and the most promising PD marker, α-synuclein.

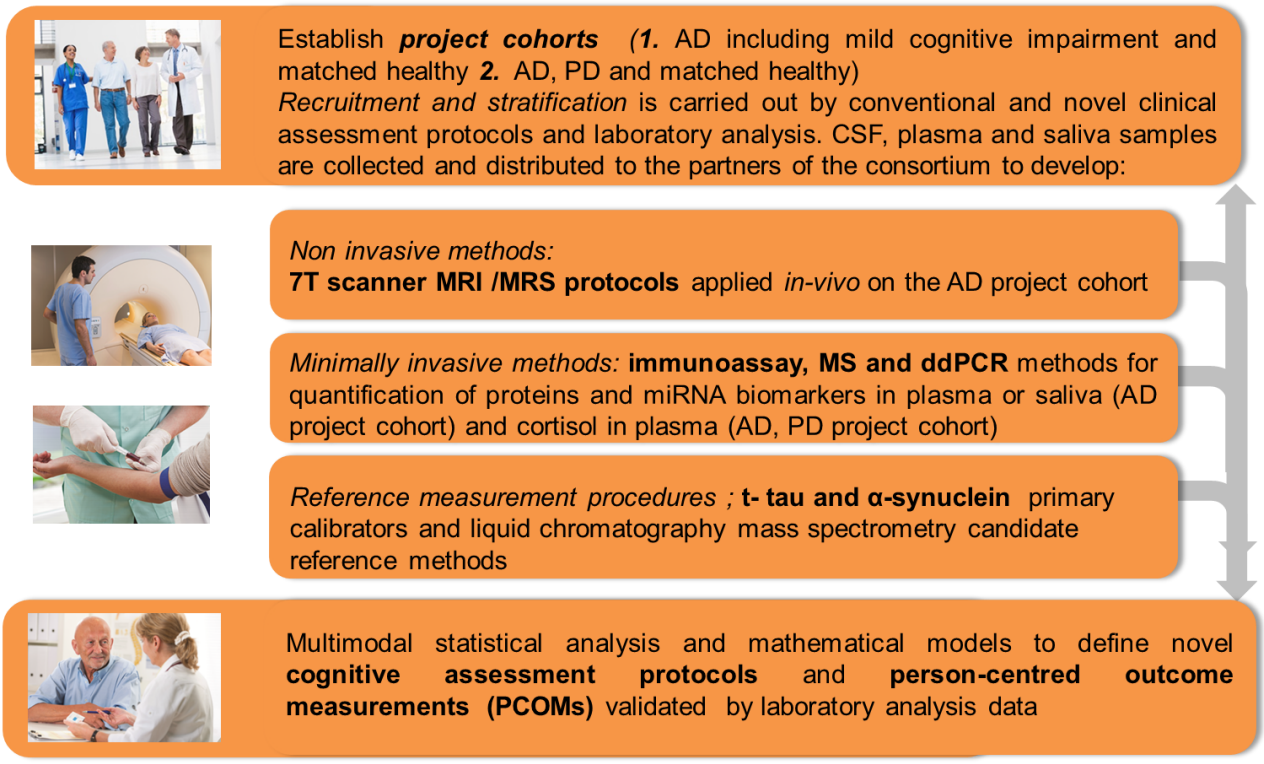
5. **To determine and characterise patient-centred outcome measures (PCOM) for NDD**. This will be achieved by developing improved clinical assessment questionnaires focused on the decline in motor and cognitive functions and increases in behavioural, communicative and psychological symptoms (e.g. agitation) in AD patients. For PD patients, this will be also supported by the validation of a novel immunoassay for the measurement of the stress biomarker cortisol, with the ultimate aim beyond this project of using this immunoassay to improve the assessment of the efficacy of therapeutic interventions administered to reduce anxiety.
6. **To develop, validate and verify multimodal statistical analyses** used to correlate NDD patient health status with AD biomarker and MRI data. These novel analyses will identify the most promising tools for NDD early diagnosis and disease progression monitoring, aiming for better resolution in detecting the early signs of disease and reducing the number of biomarker studies required (e.g. in cases where biomarker studies are especially challenging).
7. **To form a NeuroMet Stakeholder Network** in Europe for neurodegenerative disease diagnosis and disease progression monitoring. The NeuroMet Stakeholder Network will include the consortium (and NMI partners) and at least 10 relevant NDD stakeholders including clinicians, instrument manufacturers and national and international organisations to facilitate the uptake of the technologies and measurement infrastructure developed by the project and to ensure that patient and measurement needs are met.

4 Results

Important progress was made in NeuroMET towards the application of metrological concepts and procedures to underpin NDD early diagnosis. Particularly:

- a **metrological and multidisciplinary network of laboratories** was established providing reference measurement systems and tools to underpin the development and validation of minimally/not invasive methods for *early diagnosis* of Alzheimer's and Parkinson's diseases and *drug therapeutic monitoring* (objectives 1-6)
- **metrological concepts were applied throughout the workflow for AD diagnosis** from cognitive assessments (objective 5 and 6), to magnetic resonance imagining and spectroscopy (objective 2), and fluid biomarkers (objective 3 and 4)
- a **prototype Memory Metric** was developed that will be validated in the follow-on project 18HLT09 NeuroMET2 for further implementation into clinics (objective 5 and 6)

The structure of the NeuroMET project is shown below.



4.1 Objective 1. To establish patient cohorts representative of AD (including MCI due to AD and dementia due to AD), PD neurodegeneration and a matched group of healthy control subjects.

Two patient cohorts were established within NeuroMET, to be used as a reference for the development and validation of metrological methods, and to generate data to enable correlation of the NeuroMET metrological results with the results from larger clinical studies.

Cohort 1: participants were recruited by Charité in Berlin. Men and women between 55 and 90 years of age with normal vision with or without aid and with ability to consent were included. Exclusion criteria were stroke, Morbus Parkinson, untreated or severe depressive episodes, newly initiated therapy with Acetylcholinesterase (AChE) inhibitors/memantine, pregnancy, other neurological disorders, history of drug or alcohol abuse, and non-suitability for MRI (e.g. persons with claustrophobia, active implants or ferromagnetic implants such as pacemakers). In total 39 Healthy Controls (HC), 23 MCI patients and 26 AD patients were recruited with an average age of 70.8, 70.7, and 75.3, and education years 14.7, 15.3, and 13.9, respectively for the three patient groups. All participants had a first baseline examination – divided into two visits - where a battery of psychometric tests was carried out.

During the first participant visit blood, saliva, and CSF samples (when applicable) were collected. Usually within the same week, participants came in for the second visit, also in order to be scanned at PTB using the 7 Tesla scanner (part of objective 2). Furthermore, all participants underwent standard neurological and routine examination at Charité, that included standard blood tests to rule out a number of different diseases such as infections, anemia, leukemia. When applicable standard CSF measurements related to AD (Aβ-40, Aβ-42, total-tau (t-tau) and phosphorylated-tau (p-tau)) were also carried out. 66 participants also attended a follow up appointment where they underwent the same battery of psychometric tests as at baseline visit, neurological examination and again saliva samples were collected.

The test battery included 11 questionnaires (e.g. Clinical Dementia rating and Day Sleepiness Questionnaire) and 12 neuropsycological tests (e.g. Digit Span Test and the Auditory Verbal Learning Test). The cognitive assessment data from the baseline and follow up visits were sent to RISE for the development of validated PCOMs and mathematical models in objective 5 and 6. The collected NeuroMET blood, CSF, and saliva samples were sent to CHU Mpt, LGC and INRIM for biomarker analysis (objective 3). Blood and saliva samples were also used by LGC and INRIM to develop improved assays, and by CHU Mpt to generate data by using commercially available kits including the recently developed digital ELISA method by Simoa (also objective 3).

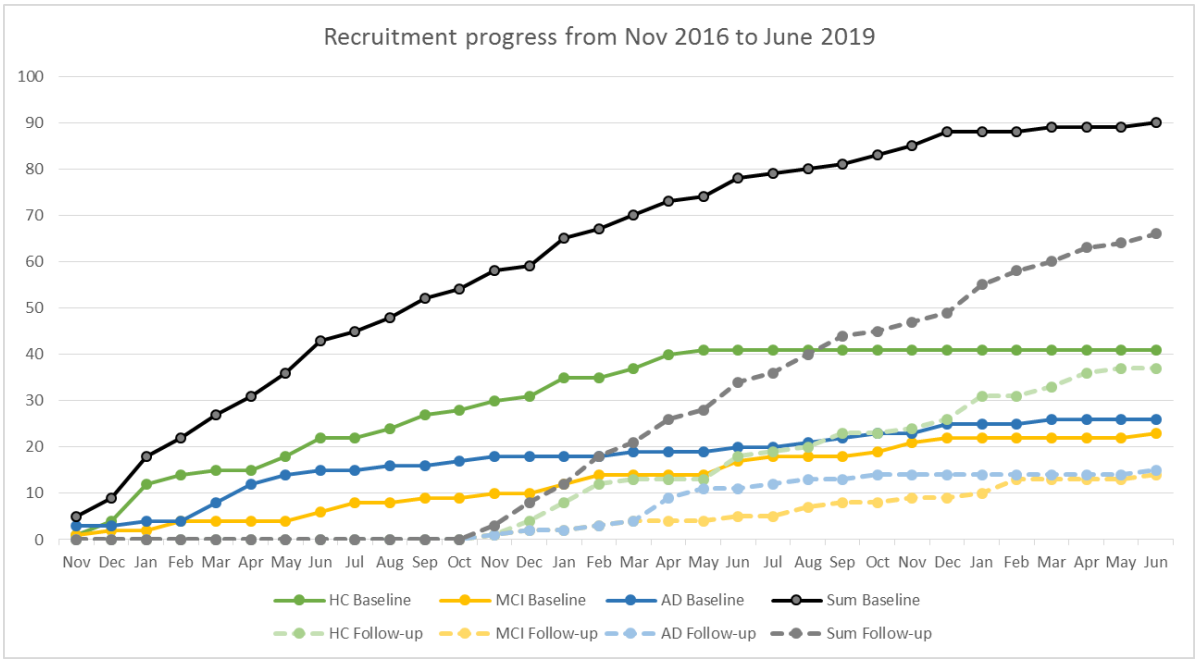


Figure 1. Cohort 1 recruitment progress/month

CSF samples were shipped to CHU Mpt to generate CSF biomarker data. The biomarkers were established biomarkers for AD and PD and served as an anchor to the data obtained from the analysis of blood and saliva.

Cohort 2: participants were recruited by UEA for a feasibility study on the potential of using cortisol measurements to monitor stress in PD patients (objective 5). Stress and anxiety are important clinical aspects for PD patients, which need to be addressed to improve their quality of life. Stress may also interfere with the efficacy of drugs.

Exclusion criteria for the UEA cohort were Addison's disease and steroid treatment. A total of 4 AD, 8 PD and 10 HC participants were recruited with average ages of 79.0, 75.6, and 70.9 respectively. All participants completed a self-reported anxiety scale questionnaire, the generic Hospital Anxiety and Depression Scale (HADS).[1] Additionally PD patients completed the Parkinson's Anxiety Scale (PAS),[2 3] and AD participants the Rating of Anxiety in Dementia (RAID).[4] Blood samples were collected by a specialised nurse at the participant's home on three separate days within a period of 2 weeks and at three time points per day (11am, 1pm and 3pm). Samples were then shipped to LGC for immunoassay analysis of free and total cortisol as marker of stress (to be used in objective 5).

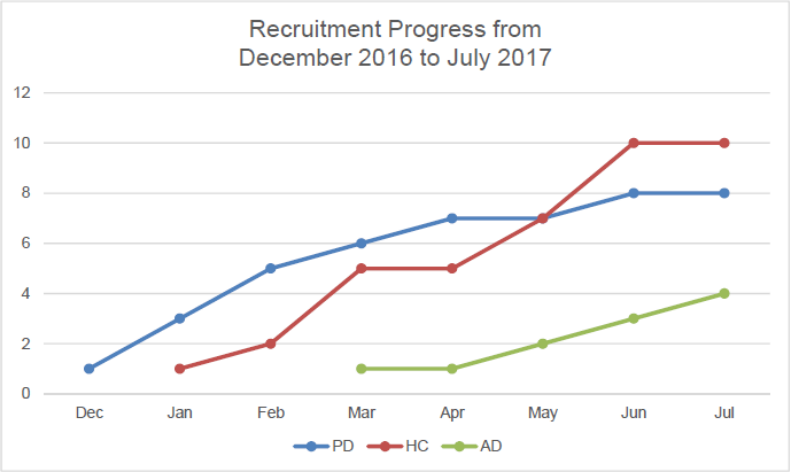


Figure 2. Cohort 2 recruitment progress/month

Summary

Two patient cohorts were established: cohort 1 (Healthy matched Controls (HC), Mild Cognitive Impairment (MCI), and AD) was used to develop validated PCOMs for AD and MCI patients through MRI and magnetic resonance spectroscopy (MRS), protein, microRNA and statistical analysis; cohort 2 (HC, AD and PD) was used to assess cortisol as a potential marker for stress in PD and AD patients.

MRI and MRS on a 7 T scanner were carried out on all patients from cohort 1, and CSF and blood samples were distributed to the project partners for immunoassay and ddPCR analysis of biomarkers. The cognitive assessment data from cohort 1, the MRI, MRS (objective 2) and fluid biomarker data (objective 3) were also collated for statistical analysis and the development of Construct Specification Equations (CSE) in objective 6. In addition, blood samples from patients from cohort 2 were analysed by immunoassay for free and total cortisol as markers of stress (objective 5). The data from objective 1 on cognitive assessments were then used in objectives 5 and 6 to input into the development of a Memory Metric (objective 5) and in the mathematical model (objective 6).

The objective was achieved and the maintenance of cohort 1 will continue in the follow-on project 18HLT09 NeuroMET2 for longitudinal studies.

4.2 Objective 2. To develop and validate non-invasive MRI approaches for in vivo characterisation of AD and MCI patients and healthy matched controls.

Neuroimaging techniques are very promising non-invasive approaches for early detection of NDD and can be used to improve understanding of underlying principles of neurodegeneration. Neuroimaging techniques are also currently used in routine practice for diagnosis of AD. Hippocampal atrophy determined by MRI, together with Aβ42 in CSF, is one of the two biomarkers approved by the European Medical Agency for clinical trials. In this project, MRI protocols were developed on an ultrahigh field scanner (7 T) to generate accurate data to underpin the validation of PCOMs in objective 5. Furthermore, for the first time, quantification of neuro-metabolites in AD was attempted at 7 T, and the *in-vitro* measurement uncertainty was determined for the applied measurement sequence combined with the used measurement equipment. Finally, a 3 T protocol was derived to enable the acquisition of similar measurands at a more clinically accessible field strength.

All *in-vivo* MR measurements were performed at PTB using a 7 T whole body Magnetom MR system (Siemens Healthineers, Erlangen, Germany) equipped with a 1-channel transmit/32-channel receive head coil (Nova Medical Inc., Wilmington, MA, USA). Each *in-vivo* MR measurement on the participants from cohort 1 consisted of three major sections:

1. structural or “morphometric” measurements, in order to assess structural information on each individual participant’s brain structure and anatomy;
2. MR spectroscopy (MRS) measurements, to measure concentrations of neurometabolites within the brain tissue of participants; and
3. resting-state functional MRI (rs-fMRI), to extract information on functional connectivity within the participant’s brains.

After analysis, the obtained results were sent to RISE to validate PCOMs and to develop mathematical models to be used towards early diagnosis of AD and its uncertainty (as part of objectives 5 and 6).

Morphometric Measurements

T1-weighted images were acquired to obtain structural information about the participants’ brains from cohort 1 and reveal structural changes as potential AD biomarkers. This was achieved by using an MP2RAGE[5] [6] sequence (TR/TE = 5000 ms/2.51 ms; TI1/TI2 = 5000 ms/900 ms/2700 ms; α1/α2 = 7°/5°; resolution: 0.75 × 0.75 × 0.75 mm^3; 2-fold acceleration) with denoised reconstruction. Typical examples of such T1-weighted images are reported in Figure 3.

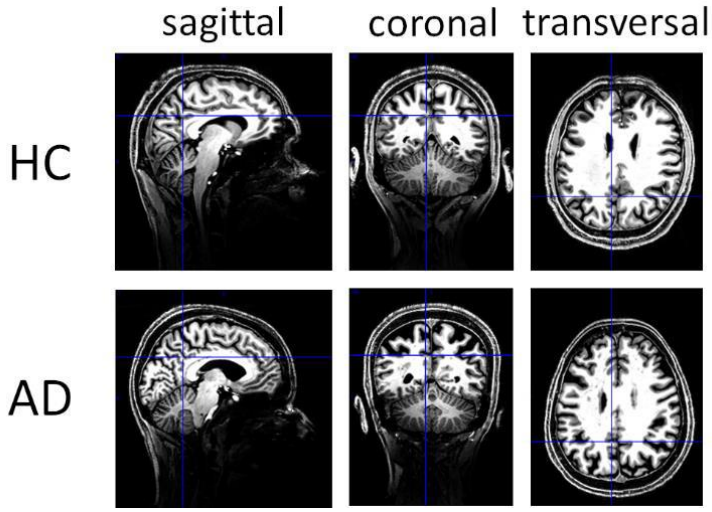


Figure 3: MP2RAGE images (TR/ TE = 5000 ms/2.51 ms; TI1/TI2 = 5000 ms/900 ms/2700 ms; α1/α2 = 7°/5°; resolution: 0.75 × 0.75 × 0.75 mm^3; 2-fold acceleration) of a HC (top panel) and an AD (bottom panel) from cohort 1 acquired at 7 T.

The T1-weighted images were then segmented into grey matter (GM), white matter (WM) and CSF, using SPM12 (Statistical Parametric Mapping 12, Wellcome Trust Center, London, UK) – a MATLAB (Mathworks Inc., Natic, USA)-based toolbox – and CAT12 (Computational Anatomy Toolbox, Structural Brain Mapping Group, University of Jena, Germany)[7] – an extension of SPM12 specifically for morphometric applications. The segmented images were used to calculate the volumes within the cranium occupied by GM, WM and CSF. Figure 4 shows an example of segmentation and in Figure 5 the segmented images are reported for a HC and AD participant (cohort 1). It is possible to observe in Figure 5 the increased CSF volume occupied in the AD patients which is a direct consequence of brain atrophy due to AD.

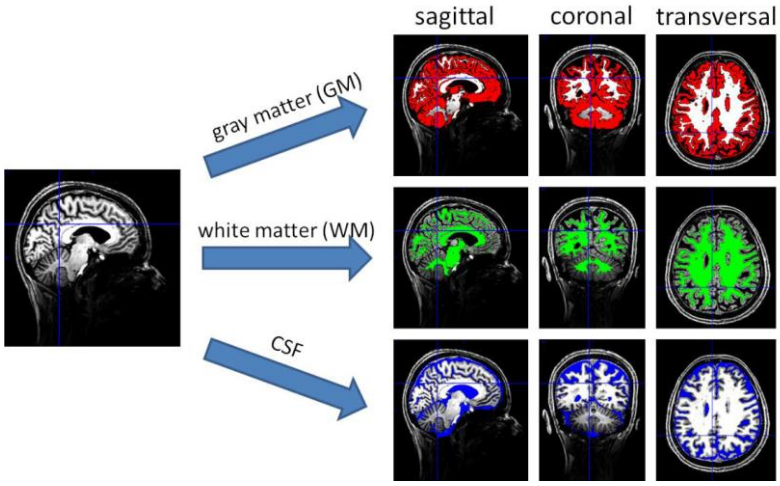


Figure 4. Anatomical Images are being automatically segmented into GM, WM, and CSF, and the respective volumes are calculated.

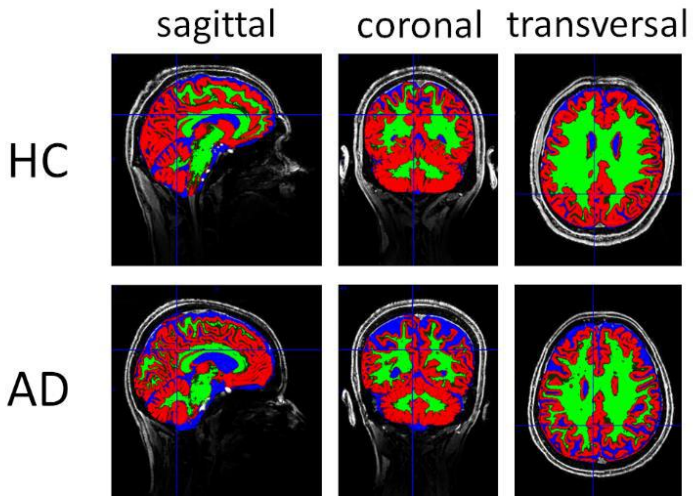


Figure 5. Segmented images of a HC and an AD participant from cohort 1. It can be seen, that the AD patient's brain is more atrophic, i.e. the GM fraction (red) is decreased and the volume occupied by CSF (blue) is increased compared to the HC

The GM was further segmented into cortical and subcortical structures by using open source or GNU General Public Licensed softwares such as CAT12[7] or FreeSurfer[8] (Laboratory for Computational Neuroimaging, MGH, Harvard, MA, USA). The calculated volumes of cortical and subcortical structures or regions of interest (ROIs) were then normalised to the total intracranial volume (TIV).[9] [10]

The volumes of several ROIs, known to be affected in AD, were compared between the different groups within the cohort (i.e. HC, MCI and AD) and statistically significant differences between groups were observed for example for the total grey matter volume (GM volume), the mean thickness of the cortex, the left and right anterior lateral temporal lobe (lAntLatTeLo and rAntLatTeLo), the left and right hippocampus (lHC and rHC), the left and right inferior lateral parietal lobe (lInfLatPaLo and rInfLatPaLo), the left and right posterior cingulate gyrus (lPosCinGy and rPosCinGy), and the left and right putamen (lPut and rPut).

In order to estimate the ability of the GM volume and the cortical thickness to correctly classify subjects as HC, MCI or AD, the receiver operating characteristics (ROCs) and the areas under the ROC curves (AUCs) were calculated (Figure 6).

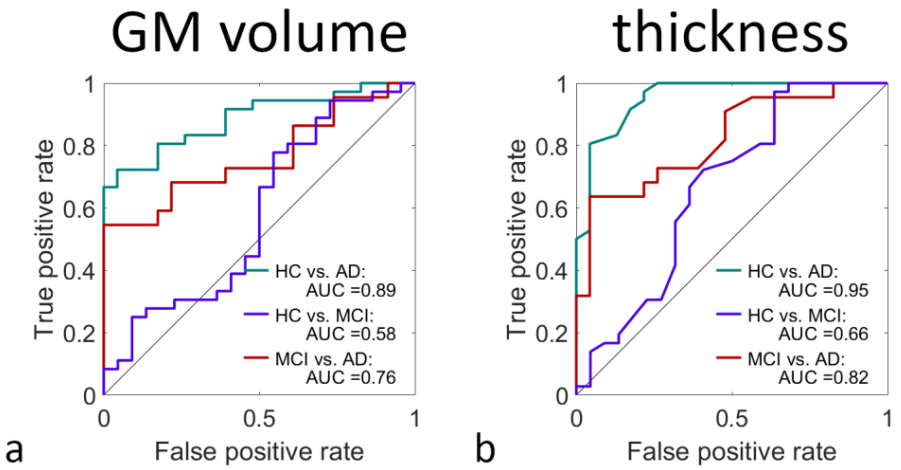
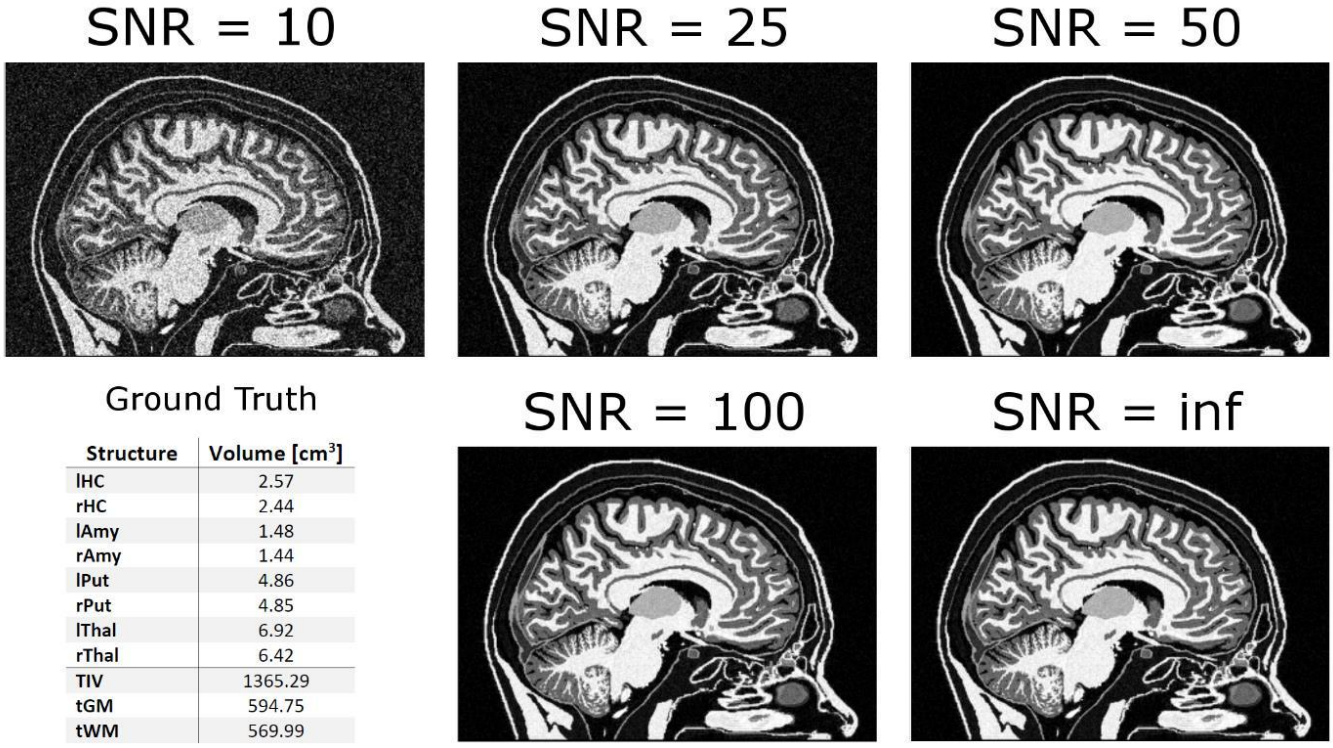


Figure 6. ROCs for the GM volume and cortical thickness as differentiator between: HC and AD patients in green; HC) and MCI patients in purple; and MCI patients and AD patients in red. The area under the curve was calculated for each ROC and is given in each plot.

Uncertainty

In order to compare robustness and reliability, and to estimate measurement uncertainties for the calculated ROI volumes, a virtual phantom was developed and used to simulate MRI data. In a first attempt, each of the 153 brain structures within the MIDA voxel model[11] was assigned an empirical signal intensity value, based on average observed image intensities from anatomical data acquired in cohort 1. Figure 7 displays the same slice of the simulated data sets at different signal to noise ratios (SNR).



Structure	Volume [cm³]
lHC	2.57
rHC	2.44
lAmy	1.48
rAmy	1.44
lPut	4.86
rPut	4.85
lThal	6.92
rThal	6.42
TIV	1365.29
tGM	594.75
tWM	569.99

Figure 7. Slice of the simulated data sets at five different levels of signal-to-noise ratio (SNR)

Spectroscopic Measurements

The MR spectroscopy voxel (20 × 20 × 20 mm^3) was positioned in the posterior cingulate cortex (PCC, see inlay in Figure 10). Magnetic resonance spectroscopy (MRS) was then performed using a SPECIAL[12] [13] localisation sequence with a very short echo time (TE = 9 ms). The concentrations of the following markers were calculated: aspartate (Asp), choline (Cho), creatine (Cr), γ-amino butyric acid (GABA), glutamine (Gln), glutamate (Glu), glutathione (GSH), myo-Inositol (Ins), lactate (Lac), N-acetyl aspartate (NAA), N-acetyl-aspartylglutamate (NAAG), scyllo-inositol (Scyllo), and taurine (Tau). Those concentrations were corrected for CSF fraction within the voxel, obtained from the segmentation of the T1-weighted images, and relaxation. Examples for typical high-quality spectra acquired in a HC subject and an AD patient are shown in Figure 8.

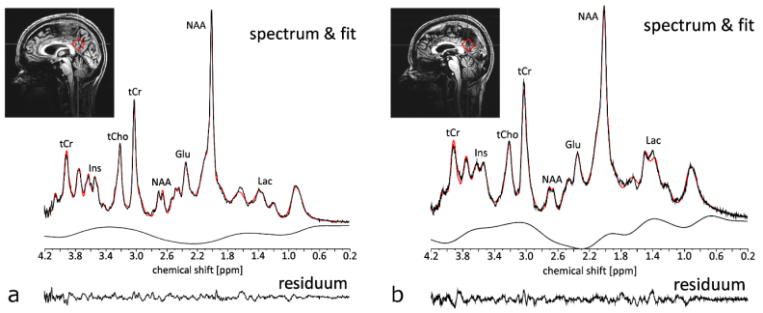


Figure 8. Typical MR spectrum acquired in the posterior cingulate cortex (PCC) from a HC (a) and an AD patient (b) from the NeuroMET cohort 1. The inlays on the left to the spectra depict the respective voxel positions. The spectral data (black) are overlayed with the LCModel fit (red). The bottom panel displays the residuum.

Data analysis focused on the biomarkers NAA, Ins and Cho were reported to be altered in AD in previous studies[14], as well as biomarkers Glu and GABA, which were suspected to be altered in AD.[15] Figure 9 shows the AUCs of the 5 considered biomarkers, as obtained from measurements from cohort 1:

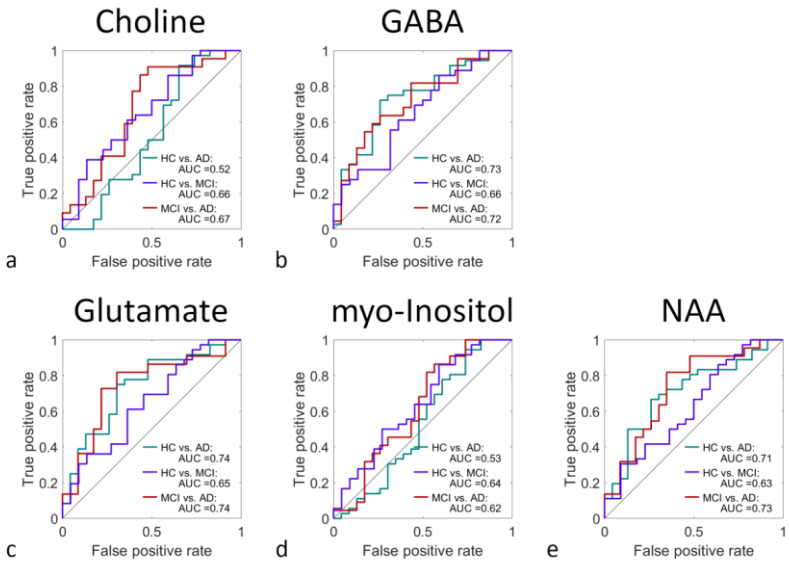


Figure 9. Receiver Operating Characteristics (ROCs) for concentrations of choline, GABA, glutamate, myo-inositol, and NAA as differentiator between: 1) HC and AD patients in green; HC and MCI patients in purple; and MCI patients and AD patients in red. The AUC was calculated for each ROC and is given in each plot.

The most promising MR-derived biomarkers for AD, identified within this project, underpin some of the findings from earlier studies. These include the GM volume adjusted for the total intracranial volume, the average cortical thickness of the brain, the volumes of the lHC and rHC, and the lPosCinGy and rPosCinGy. Furthermore, the volumes of the lPut and rPut appear to differentiate between HC and MCI subjects with increased accuracy compared to other volumes of subcortical structures.

The results from earlier MRS studies, claiming an alteration of Ins over the course of the progression of AD could not be confirmed within the results of this project. However, alterations were observed, in the concentrations of NAA, GABA and Glu.

Uncertainty

During this project, a phantom spectroscopy measurement was performed each month in order to ensure high data quality and proper functionality of the 7 T scanner, and to gather an initial estimate of measurement uncertainty in *in vitro* MRS measurements. The mean concentrations of the metabolites Cho, Cr, GABA, Gln, Glu, Ins, and NAA, which were contained in the used phantom, were calculated together with the estimated standard deviations, and the estimated standard uncertainties[16] of these metabolites (table 1). Figure 10 displays the difference from the mean concentration of each metabolite in percent, measured in individual monthly measurements. However, it has to be noted that sources of measurement uncertainties of *in-vivo* applications are higher in number and more complex to assess.

Table 1: Mean metabolite concentrations, estimated standard deviation from mean, estimated standard uncertainty, and mean CRLBs of monthly phantom measurements.

Metabolite	mean concentration [mmol/l]	estimated standard deviation (s) [mmol/l]	estimated standard uncertainty (u) [mmol/l]	estimated standard uncertainty (u) [%]	mean Cramer-Rao Lower Bound [%]
Cho	1.328	0.045	0.008	0.6	3.6
Cr	4.587	0.329	0.058	1.3	3.1
GABA	3.089	0.328	0.058	1.9	8.0
Gln	2.041	0.087	0.015	0.8	13.2
Glu	9.088	0.594	0.105	1.2	3.1
Ins	4.975	0.179	0.032	0.6	4.1
NAA	11.828	0.555	0.098	0.8	2.0

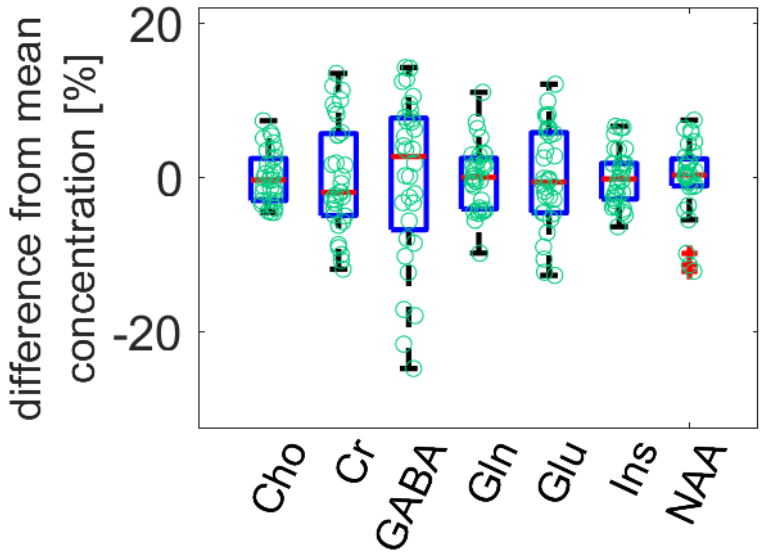


Figure 10: Difference from mean concentration of metabolites in percent. The boxes extend from the 25th to the 75th percentiles, including the median value, which is indicated by the central mark. Outliers are plotted individually (+). Each green circle represents an individual measurement.

Functional Measurements

A multi-band Echo-Planar Imaging (EPI) sequence[17] (TE/TR = 28.2 ms/2200 ms; 1.5 mm isotropic resolution) was used to acquire 300 whole brain images over a time-course of 11 minutes. fMRI analysis is computationally expensive and subject motion is detrimental to data quality and reliability of calculated functional connectivity results. An initial analysis of fMRI measurements was performed on a subset of 19 HC subjects, 16 MCI patients and 10 AD patients from cohort 1. This was due to the costs of the experiments and the limitations on data quality generated by motion. Signal fluctuations and the comparison of such fluctuations between different ROIs were used to calculate resting-state functional connectivity ROI-to-ROI resting-state functional connectivity between the PCC and the lHC and rHC. It was noted that for this subset of samples functional connectivity between lHC and PCC was positively correlated to GABA concentration. Therefore, a more extensive analysis including the whole cohort is needed, where data quality is not impaired by subject motion.

Summary

MRI and MRS scan parameters were optimised on a 7 T MR system and protocols were established to be applied *in-vivo* on the patients from cohort 1. Each *in vivo* MR measurement consisted of three major sections: 1) structural or “morphometric” measurements, in order to assess structural information on each individual participant’s brain structure and anatomy; 2) MRS measurements, to measure concentrations of neurometabolites within the brain tissue of participants; and 3) resting-state functional MRI, to extract information on functional connectivity within the participant’s brains.

The most promising MR-derived biomarkers for AD identified within this project underpin the findings from earlier studies on AD. These include the GM volume adjusted for the total intracranial volume, the average cortical thickness of the brain, the volumes of the left and right hippocampus, and the left and right posterior cingulate gyrus. Furthermore, the volumes of the left and right putamen appear to differentiate between HC and MCI subjects with increased accuracy compared to other volumes of subcortical structures. Although the project’s findings supported some of those from earlier studies, the results from earlier MRS studies, claiming an alteration of myo-inositol over the course of the progression of AD could not be confirmed within the NeuroMET project. However, alterations were observed in the concentrations of N- acetyl-aspartate, (as per other studies), and in concentrations of γ-amino butyric acid and glutamate. Further work on these findings will be carried out through longitudinal studies in the follow-on project 18HLT09 NeuroMET2.

4.3 Objective 3. To develop minimally invasive methods for early diagnosis and drug therapeutic monitoring. This will be achieved by developing and applying conventional and innovative methods for the quantification of recognised and emerging biomarkers present in of the AD and PD patient cohorts.

While CSF biomarkers such as t-tau, Aβ42, and Aβ40 are recognised biomarkers for accurate diagnosis of AD, lumbar puncture is not perceived to be a minimal invasive method to be used within screening programs. Hence much effort has been put into identifying reliable biomarkers and methods for their quantification in plasma.

This project addressed this by generating biomarker data in plasma and CSF from cohort 1 using commercially available assays and improved assays. All data generated was sent to RISE for validation of PCOMs and the development of mathematical models to underpin AD diagnosis in objective 5 and 6.

A number of biomarkers were selected in consultation with stakeholders to:

- generate biomarker data in plasma, CSF, and saliva from cohort 1 for objective 5 and 6 by using commercially available methods.
- develop novel assays to increase sensitivity and accuracy of NDD biomarkers in plasma and saliva when commercial kits are not available.

The biomarkers selected were: Aβ42, Aβ40, t-tau, neurofilament light (NFL), α-synuclein, and cortisol.

Aβ42, Aβ40, t-tau, the ratio Aβ42/Aβ40, and t-tau in CSF are recognised AD biomarkers. Furthermore Aβ42, Aβ40, and NFL have shown great potential over the past few years as biomarkers for AD in blood. The introduction to the market of digital ELISA and MS methods for Aβ42 and Aβ40 is rapidly accelerating the uptake of these biomarkers, but efforts are still required to address their measurement standardisation, definition of thresholds, and pre-analytical issues.

CHU Mpt analysed the CSF, blood, and saliva samples from cohort 1 sent by partner Charité for Aβ42, Aβ40, t-tau, NFL, and α-synuclein using Meso Scale Discovery (MSD) or SIMOA. All data generated was sent to RISE for validation of PCOMs, development of mathematical models for early diagnosis of AD and definition of its uncertainty in objective 5 and 6.

LGC developed a new approach based on immunoassay and standard addition to increase sensitivity and reduce matrix effects when measuring biomarkers in e.g. plasma. The assay and method were developed for use with Aβ42 and Aβ40, and then validated using certified reference materials. The validated method was then applied to the blood samples from cohort 1. Additionally, LGC measured free and total cortisol as markers of stress and as a potential NDD biomarker in the blood samples of both cohort 2 (objective 5) and cohort 1 respectively. The data from cohort 1 was sent to RISE to be used in objective 5 and 6.

INRIM and LGC selected three miRNA markers (Let-7g-5p, hsa-miR-15a-5p, hsa-miR-34a) to be potentially measured in cohort 1 by digital polymerase chain reaction (dPCR) and immunoassay. Early experiments showed higher sensitivity in dPCR experiments over immunoassays and therefore the ddPCR approach was then applied to the samples from cohort 1.

Finally, LGC explored the potential use of the LC-MS method for quantification of α-synuclein in CSF (objective 4) to quantify α-synuclein in saliva. However, the method did not show enough sensitivity to be applied to clinical samples.

Amyloid beta (Aβ) peptide detection in blood-based samples by standard additions

A new method was developed to enable the detection of low-level biomarkers of interest (such as Aβ peptides) in blood samples (i.e. plasma and serum) to overcome issues due to complex matrix interferences that often plague conventional immunoassays. The method developed in the project is based on a standard addition approach that is traditionally used for chemical analysis but was adapted for use with immunoassays. With chemical analysis, standards are spiked directly into the test sample and the signal response is measured, which gives rise to a linear correlation between the signal and the concentration of the spike. When this approach was implemented for immunoassays the correlation between the signal response and the total (i.e. the spiked and endogenous) concentration of the analyte was sigmoidal on a log-log plot. However, as the central portion of the sigmoid produced from the log concentration-log signal plot essentially appears linear, the logarithm of signal outputs corresponding to the logarithm of the total concentrations within the linear

portion of the curve may be used to derive the unknown endogenous analyte concentration. The endogenous analyte concentration was therefore ascertained by using the minimum residual sum of squares as the trendline of maximum linearity will exhibit the smallest residual sum of squares. The full method is described in Pang et al.[18] and a schematic representation of the approach is presented in Figure 11.

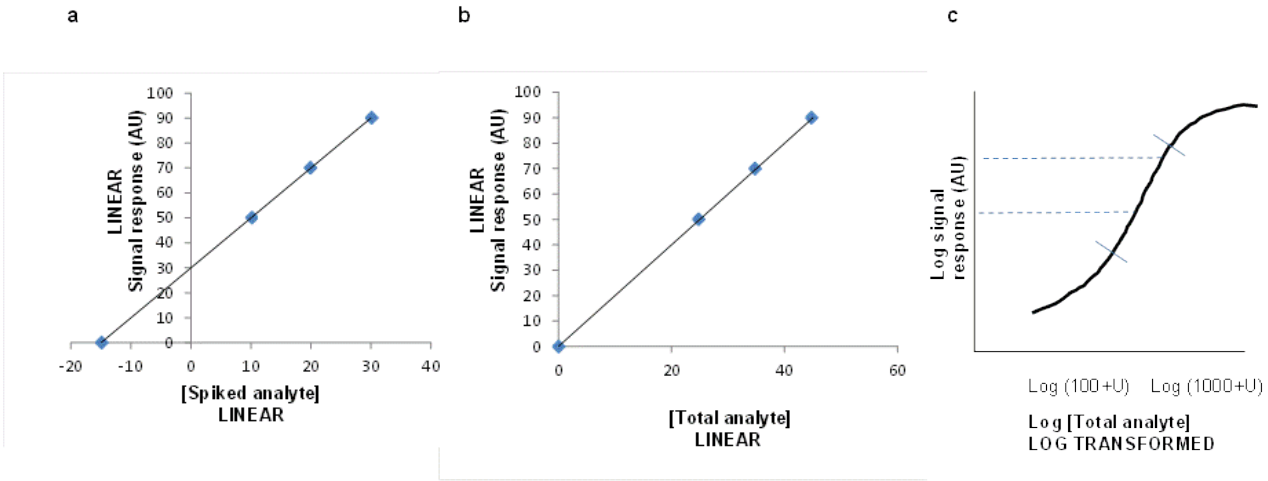


Figure 11: Standard additions for chemical analytes and biological targets by immunoassay. Fig. 10a: A typical plot for illustrating the method of standard additions for chemical analysis. The intercept at -15 indicates the endogenous concentration of the chemical analyte in this example is 15 units. Fig. 10b: An alternative view of Supplementary Fig 10c to show the relationship if the total analyte (not just spike concentration) is plotted on the x-axis, where U denotes the unknown endogenous chemical analyte concentration. If a linear change in the concentration of a chemical analyte exhibits a linear change in response, the gradient can be resolved for the line of fit, and the unknown endogenous concentration can then be determined; i.e. using the equation for a straight line: y= mx +c, where y is the response, x= concentration and c= intercept. Therefore, with c=0, 70-50=m(20+U-10-U) gives a gradient of m= 2. Therefore, if y= 2x, then 50= 2(10+U), and U=15 units. Fig. 1c: The analogous plot for the signal response curve for an immunoassay of a biological analyte where part of the log transformed response and log transformed total analyte concentration give rise to a linear correlation which cannot be solved in the same manner as shown as for a chemical analyte (as shown in Fig 1b) in spite of a linear correlation within the sigmoid, due to the log transformation of the concentrations on the x-axis.

For the analysis of Aβ40 and Aβ42 peptides within human plasma, the MSD Aβ Peptide Panel 1 (4G8) kit was used with the difference that the kit calibrants were used to spike the test samples directly. By spiking individual test samples (typically eight) with distinct net spike concentrations, the endogenous analyte of the Aβ peptide was then determined using a solver with the statistical program R, whereby the known net spike concentrations and the corresponding signal were inputted for each of the spiked samples.

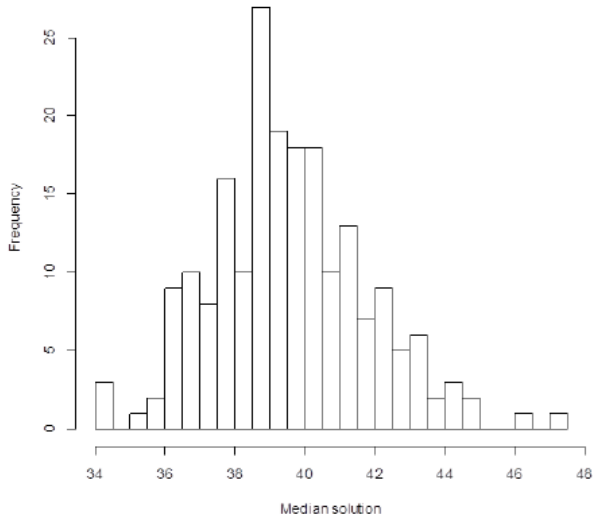


Figure 12. Histogram depicting the Monte Carlo simulation for the Aβ40 assay performed for Donor 83 (exhibiting an uncertainty estimate of 3.65 %)

The average standard uncertainty estimates for the Aβ40 and Aβ42 assays using the standard additions approach were 9.8 % and 5.9 % respectively. The coverage factor is 2 and the estimates were based on Monte

Carlo simulations (with n=200). The Monte Carlo method was used to estimate the contribution of sampling error in the signal response (measurement repeatability) and was carried out by using the observed residual standard deviation as the sampling distribution and repeating the procedure 200 times (Figure 12). The final estimate and uncertainty were obtained from the distribution of simulated results.

Aβ40 and Aβ42 measurements were carried out on all blood samples from cohort 1 and the data was sent to RISE for validation of PCOM in objective 5 and data analysis in objective 6.

It should be noted that while the standard additions approach method has been developed and applied for the analysis of amyloid β peptides, the approach is generic and can only currently be applied across a broad range of analytes and matrices.

Cortisol immunoassay measurements

Cortisol is a steroid that plays a vital role in the regulation of essential physiological processes and is expressed according to circadian rhythms. Raised levels of cortisol are a well-recognised endocrine sign of anxiety, depression, and sleep disturbances, and are also raised in neurological disorders such as PD.

Measurement of anxiety levels is problematic in both PD and AD given the subjective nature of self-reporting. Therefore, as cortisol is raised with anxiety, it may be a helpful objective surrogate measure for anxiety in AD and PD. Hence a pilot study was undertaken on cohort 2 (objective 5) to assess if there was any utility in gauging anxiety levels by measuring cortisol in serum derived from blood. Cortisol was also investigated as a putative marker for NDD in cohort 1 samples. This was based on stakeholder suggestions and literature citing cortisol as a potential NDD biomarker.

Traditionally it has not been possible to measure free cortisol in serum directly. However, a method was established to ascertain the direct measurement of both free and total cortisol in serum without the need to fractionate the free portion of cortisol from the serum using ultracentrifugation, ultrafiltration, or extensive dialysis.

The MESO QuickPlex SQ120 was used to measure both free and total cortisol in serum samples from cohort 1 (as a NDD biomarker) and cohort 2 (as a marker of stress). A schematic representation of the assay used for cortisol measurements is in Figure 13.

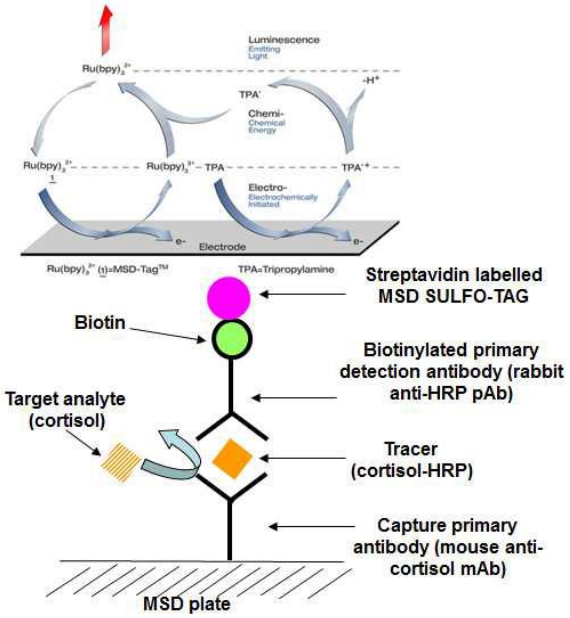


Figure 13. Schematic representation of the MSD immunoassay for cortisol detection

Free and total cortisol were measured in the samples from cohort 1 as potential markers of AD. The variability in both the free cortisol and the total cortisol measurements were due to sourcing the blood on different times and days and was able to be minimised by deriving the percentage of free cortisol.

The function of the assays developed were validated using certified reference materials from NIST (NIST cortisol 921 and NIST SRM 971: hormones in serum). Recoveries of total cortisol from the male NIST sera ranged from 110.6 – 122.6 %. The uncertainty of the total cortisol assay was 11.8 % and the uncertainty of the free cortisol assay was 13.9 %.

The results obtained for cohort 1 were sent to RISE for data analysis as part of objective 6.

Development of novel minimally invasive methods for quantification of AD biomarkers – droplet digital polymerase chain reaction (ddPCR) methods

miRNAs are a class of small non-coding RNAs, protected by membranes in micro-vesicles, that are detectable in bodily fluids. More than 1500 human miRNAs have been discovered and several miRNAs have been recognised and proposed as AD biomarkers. The miRNA expression pattern in an AD patient's brain is altered and it has been reported in the literature that some miRNAs are downregulated (e.g. those involved in the formation of β-amyloid plaques, accumulating outside nerve cells and becoming toxic to neurons in the brain), while others are upregulated (e.g. those with a possible role in neuronal inflammation in AD). Based on recent discoveries, quantification of miRNAs in body fluids has been proposed as a new tool for AD diagnosis. However, although diverse approaches have been implemented to analyse the presence of miRNA in bodily fluids, current detection methods lack standardisation.

INRIM developed a ddPCR method for detection and accurate quantification of candidate AD miRNA biomarkers. miR-26a-5p, which was selected as good candidate for a possible comparison with immunoassay method and was analysed on 27 blood samples from cohort 1 and miR-34a (middle-late stage AD, upregulated in AD) were analysed on all the samples from cohort 1 and miR-34a (middle-late stage AD, upregulated in AD) were analysed on all the samples from cohort 1.

ddPCR has a much higher sensitivity than that of traditional real time PCR and can detect low copy numbers of DNA without the need for external standards. This leads to a reduction in the uncertainty of the measurements. However critical factors affecting the reproducibility and accuracy of this technology still exist and it is known that pre-analytics play an important role in measurement variability.

Blood and CSF were collected at Charité following a protocol provided by INRIM (enabling blood stabilisation), frozen, and shipped to INRIM for analysis. miRNAs were successfully extracted from blood using the Maxwell RSC miRNA Tissue Kit from Promega, adapted for blood and CSF samples. The RNA extracted with the developed procedure was in the range of 160 to 2220 ng/uL. Its purity and absence of contaminants (expressed as 260/230 and 260/280 nm ratios) had mean values of 1.8 (with a CV of 16 %) and 1.9 (with a CV of 2 %) respectively. CV = coefficient of variation or relative standard deviation.

Uncertainty

Absolute quantification of the selected miRNAs was performed by amplification through ddPCR according to a mathematical model developed by the project. The mathematical model developed compares a selected miRNA's expression in a patient with the same miRNA's expression in a control (i.e. health donor), as follows:

$$\Delta miRNA\ exp. = |miRNA\ exp.Control - miRNA\ exp.Patient| \equiv$$
$$\equiv \left|\left(\frac{DNA\ copy\ number}{\mu L}\right) Control - \left(\frac{DNA\ copy\ number}{\mu L}\right) Patient\right| =$$
$$= \left|\frac{\sum_{i=1}^{n}\left(DNA\ copy\frac{number}{\mu L}\right) Control}{n} - \left(DNA\ copy\frac{number}{\mu L}\right) Patient\right| =$$
$$= \left|\left[\frac{\sum_{i=1}^{n}[D_f \cdot \frac{-\ln\left(\frac{N_{neg}}{N}\right)}{V_d}]}{n}\right] - \left[D_f \cdot \frac{-\ln\left(\frac{N_{neg}}{N}\right)}{V_d}\right]\right|$$

Equation 1

In Equation 1, N_{neg} is the number of negative droplets, N is the total number of droplets, V_d is the droplet volume and D_f is the dilution factor applied to the DNA sample template. The appropriate ratio between N_{neg}

and N is crucial in order to enhance the accuracy and to reduce the uncertainty of the measurements. A preliminary approach for definition of the uncertainty budget was also designed within the project.

Preliminary results from synthetic miRNAs were promising and showed good linearity and limits of detection (less than 5 ng DNA). Results on clinical samples from cohort 1 showed good reproducibility with uncertainties around 15 %. However, no clear correlation was observed between patients' diagnoses and the three selected miRNAs (i.e. miR-26a-5p, miR-34a, miR-146a and miR-15a-5p).

Summary

Nine proteins and miRNA biomarkers were selected by the consortium and in consultation with stakeholders as the most promising NDD biomarkers in plasma. They were Aβ40, Aβ42, t-tau, neurofilament light chain (NFL), α- synuclein, cortisol, Let-7g-5p, hsa-miR-15a-5p, hsa-miR-34a.

Aβ40, Aβ42, t-tau, NFL, cortisol and α-synuclein were measured in both CSF and plasma samples from cohort 1 and the data was used for the validation of cognitive assessments. In addition, a novel generic standard additions approach to overcome matrix effects associated with immunoassay measurements was developed and used for the detection of Aβ40 and Aβ42 peptides[1]. This method was used to measure the Aβ peptides within the plasma samples from cohort 1, alongside commercially available methods for quantifying the selected protein biomarkers in plasma and CSF samples. The data generated was used in objectives 6 to develop and validate multimodal mathematical model for correlation of health status, cognitive assessments and AD biomarkers.

A method for miRNA extraction and dPCR quantification was also developed by the NeuroMET project and used on the samples from cohort 1. Preliminary results obtained with synthetic miRNAs were promising, showing good linearity and limit of detection (less than 5 ng DNA). Results with the clinical blood samples from cohort 1 also showed a very good reproducibility and uncertainties of 15 %. But despite the promise shown with the method no clear correlation between patients' diagnoses was observed with the miRNA biomarkers investigated.

4.4 Objective 4. To improve NDD biomarker measurement comparability through the establishment of traceability chains. This will be achieved through the development of reference methods and reference materials for measurements, in CSF, of the established AD marker, tau protein (tau), and the most promising PD marker, α-synuclein.

Over the past five years a number of large studies have greatly improved understanding of CSF biomarkers in AD. Guidelines on pre-analytical sampling were produced, the Alzheimer's Association Quality program started in 2009[19,20], and the first certified reference material in the NDD area (ERM-DA482_IFCC ABETA 42 in CSF) was introduced to the market in 2017.

This project integrated with these activities by developing LC-MS methods to be potential candidate reference methods for t-tau and α-synuclein.

The t-tau work was carried out in parallel and in collaboration with the the International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) Working Group on CSF-biomarkers. LNE developed a candidate reference measurement procedure (RMP) for SI-traceable quantification of tau protein in CSF and, together with CHU Mpt, started the next steps towards recalibration of immunoassays.

α-synuclein is an intrinsically disordered protein with a high propensity for forming aggregates and is implicated in a number of NDD[21-25]. Most notably, it is the main constituent of Lewy bodies found in the brains of PD patients. It is for this reason that α-synuclein has been thought to be both a potential biomarker of NDD and therapeutic target for the treatment of PD. The quantification of α-synuclein in patient CSF has been the subject of a large body of literature and has mainly been achieved through the use of antibody-based assay such as ELISA. However, this approach has led to inconsistencies in results[26]. A recent comparison of immunoassay techniques across a number of labs demonstrated a good correlation across different immunoassay platforms but a large difference in the absolute quantity of α-synuclein that was measured[27]. Hence there is a need for an RMP for α-synuclein, particularly to support drug development for PD.

For both RMPs (for tau and α-synuclein), samples to be used as primary calibrators were sourced, characterised and SI-traceably quantified. A method for quantification of the target protein was also developed and validated. CSF samples for comparison with immunoassay data were prepared by CHU Mpt, however the analysis will be carried out in the follow-on project 18HLT09 NeuroMET2.

t-tau primary calibrator

LNE sourced two different materials as potential primary calibrators for the SI-traceable quantification of t-tau in CSF:

- recombinant Tau441 (rTau441) protein produced by recombinant DNA technology in *E.coli* by "Promise Advanced Proteomics"
- and synthetic peptide GAAPPGQK (156-163) [GAA]. This peptide is generally accepted as the only peptide to be used for quantification of t-tau due to the large number of phosphorylation sites.

Confirmation of the identity and purity of the material was performed by high resolution mass spectrometry (HRMS) and SI traceable quantification was carried out by amino acid analysis using certified reference materials as calibrators.

Different LC gradients were developed using different chromatographic columns in order to separate potential impurities. Since the concentration of the primary calibrator was determined by amino acid analysis, only amino acid containing impurities were considered, including truncated forms of the protein, peptides, and any components such as free amino acids which could compromise the amino acid analysis results.

The project showed that the purity of rTau441 was considered fit-for-purpose for the work it carried out. However additional studies should be carried out if the material is to be produced as commercial primary calibrator. In addition, the project showed that the purity of the peptide GAAPPGQK (156-163) [GAA] was also considered fit-for-purpose.

Amino Acid Analysis (AAA) was performed by gas-phase hydrolysis and isotopic dilution-mass spectrometry (ID-MS). The protein and the peptides were hydrolysed in the presence of isotopically labelled amino acids as internal standards, and the concentration of amino acids resulting from protein hydrolysis was measured by exact matching ID-MS. SI traceability was ensured through 1) optimisation of hydrolysis conditions to ensure

hydrolysis completeness and 2) calibration with primary calibrators obtained from the National Measurement Institute of Japan (NMIJ).

The amino acids; alanine, valine, isoleucine, leucine, phenylalanine, and proline were used for quantification of rTau441. Linear regression was obtained for each amino acid and protein concentration was finally calculated considering the mean of the concentrations of each amino acid. Analysis of variance by Anova test allowed the consistency of the data to be determined. This showed that the data for proline was substantially different from the data of the other amino acids. Therefore, this proline data was excluded from the data set. Potential reasons for the proline data could be the presence of a large number of prolines in the sequence (43 in total). In addition, it is known that that complete hydrolysis of proline rich regions can sometimes be difficult.

The final concentration for r-Tau441 primary calibrator was 263 ±16 µg/g

Glycine, Alanine and Proline were used for quantification of the peptide GAAPPGQK (156-163) [GAA] and the final concentration of the peptide primary calibrator was 755 ± 37 µg/g

For uncertainty measurements, a number of different parameters were considered:

- the uncertainty on the concentration of the primary calibrator
- the uncertainty associated with the weighing of the primary calibrator
- the uncertainty associated with the weighing of sample
- the uncertainty associated with the weighing of the labelled internal standard
- the uncertainties associated to the calibration curve

The parameters were combined in the following equation:

$$U_{pept}\ (k=2) = 2 \times u_{pept}$$

$$u_{pept} = \sqrt{{u_{Cech}}^2 + {u_{betw}}^2}$$

$$u_{Cech} = \frac{1}{nb\ of\ quantified\ amino\ acids} \times \sqrt{\sum {u_{Cpep\,(aai)}}^2}$$

$$u_{Cpep\,(aai)} = \frac{M_{peptide}}{M_{AA} \times Stoechiometry} \times u_{Caai}$$

With M = molecular mass of the peptide or the amino acid

$$C_{aai} = \frac{mass\ of\ the\ AA^{*}solution \times C_{AA^{*}} \times Q \times mass\ of\ the\ non\ diluted\ sample}{mass\ of\ the\ diluted\ sample \times mass\ of\ the\ diluent} + F$$

With Q = Qlin x Qcal

Qlin determined from the linearity of the calibration curve
Qcal determined from the preparation of the calibration samples

F = 0 with an uncertainty associated to the precision

$$u_{betw} = \sqrt{\frac{MS_{betw} - MS_{within}}{nb\ of\ replicates}}$$

With MS betw = mean square between groups, as determined by ANOVA

MS within = mean square within groups, as determined by ANOVA

Equation 2

t-tau LC-MS method in CSF

Two different calibration approaches (i) protein and (ii) peptide-based calibration were investigated for the quantification of t-tau in three CSF pool samples. Calibration standards were prepared gravimetrically by spiking a fixed amount of 15N-recombinant protein or peptide into a variable amount of 14N-recombinant tau or GAAPPGQK (156-163) [GAA]-peptide in artificial CSF (aCSF) matrix, consisting of diluted human serum.

The same amount of 15N-labelled protein/peptide was spiked into the CSF samples that were subjected to the same protocol as the primary calibration standards. This was based on the following steps: perchloric acid protein precipitation and supernatant recovery, a hydrophilic-lipophilic balance solid phase extraction (SPE), and finally tryptic digestion. LC-MS/MS experiments were carried out by using LC coupled to Orbitrap HRMS. MS detection was operated in parallel reaction monitoring mode, following 12 peptides and their labelled counterparts when using the rTau 441 primary calibrator. The GAAPPGQK (156-163) [GAA] peptide at m/z 363.2 and its labelled form at m/z 367.2 were used for quantification of t-tau in both approaches. The three most intense transitions per peptide were used as quantifiers. Data was analysed using Skyline and the ratio between the GAAPPGQK (156-163) [GAA] peptide and its corresponding labelled forms were obtained by considering the area ratios, as shown in Figure 14.

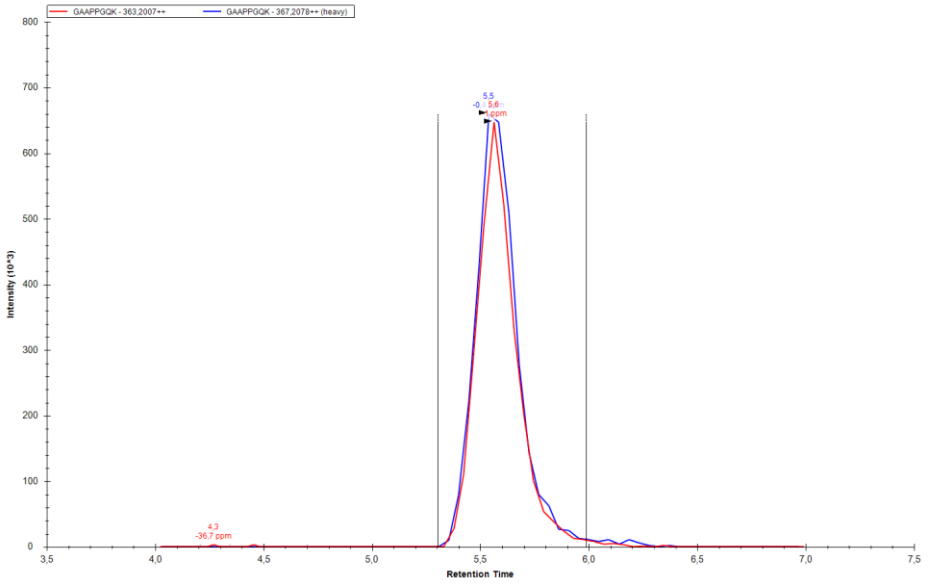


Figure 14. PRM signal of the extracted ion current of the GAAPPGQK (156-163) GAA-peptide (in red) and its15N-labelled form (in blue).

As material loss before digestion is not compensated when labelled peptides are used as internal standards, higher precision was found when rTau 441 was used as the calibrant. Therefore, the GAAPPGQK (156-163) [GAA] peptide-based calibration approach was discontinued.

The key steps which required additional investigation for the rTau441 method were (i) the homogeneity of the samples, (ii) the solvents to be used for dilution of the rTau441, and (iii) the appropriate concentrations to be used as primary calibrator. aCSF was used as solvent, but due to poor homogeneity of the primary calibrator all used aliquots were quantified by amino acid analysis.

Three pools of frozen CSF (15 mL each) were prepared by CHU Mpt and quantified for t-tau and p-tau using immunoassays. These pools were then shipped to LNE and quantified using the LC-MS method developed (see the results in table 2):

AD CSF pools	t-tau ELISA (pg/ml)	p-tau ELISA (pg/ml)	t-tau LC-MS (pg/ml)
Low	155	32	1102 ±142
Medium	470	50	2374 ±139
High	1077	135	4457 ±311

Table 2. Results of low, medium and high CSF pool samples by ELISA (Fujirebio ELISA kit) and LC-MS

The commutability of certified reference materials SI-traceably quantified using the NeuroMET LC-MS method for t-tau will be evaluated for all major immunoassays, as per the IFCC's latest recommendations in a future commutability study. Important preparations have been made for the commutability study by LNE and CHU Mpt. Three pools (>30mL) with low, medium and high values of tau were prepared by CHU Mpt and sent to LNE. In addition to this, 40 individual CSF samples of 4mL were prepared by CHU Mpt, quantified for t-tau and sent to LNE. These materials will be SI traceably quantified as part of the follow-on project 18HLT09 NeuroMET2 using the reference measurement procedure developed in this project.

In preparation for the future commutability study IVD kit manufacturers have agreed to perform measurements in either their facilities (Roche, Fujirebio, IBL) or at CHU Mpt under supervision (EuroImmun, MSD, Quanterix). The commutability study will allow the establishment of a correlation between the LC-MS method for t-tau and major immunoassays, thereby allowing an evaluation of the feasibility of recalibrating immunoassays for quantification of tau.

α-synuclein primary calibrator

α-synuclein was produced by UCL using BL21-competent cells grown in Lysogeny broth in the presence of ampicillin (100 µg/ml) at 37°C. The α-synuclein protein was purified by gel filtration and concentration was estimated from the absorbance at 280 nm using an extinction coefficient of 5600 M^{-1} cm^{-1}. Aliquots of 1 mg were lyophilised for further characterisation (see Figure 15).

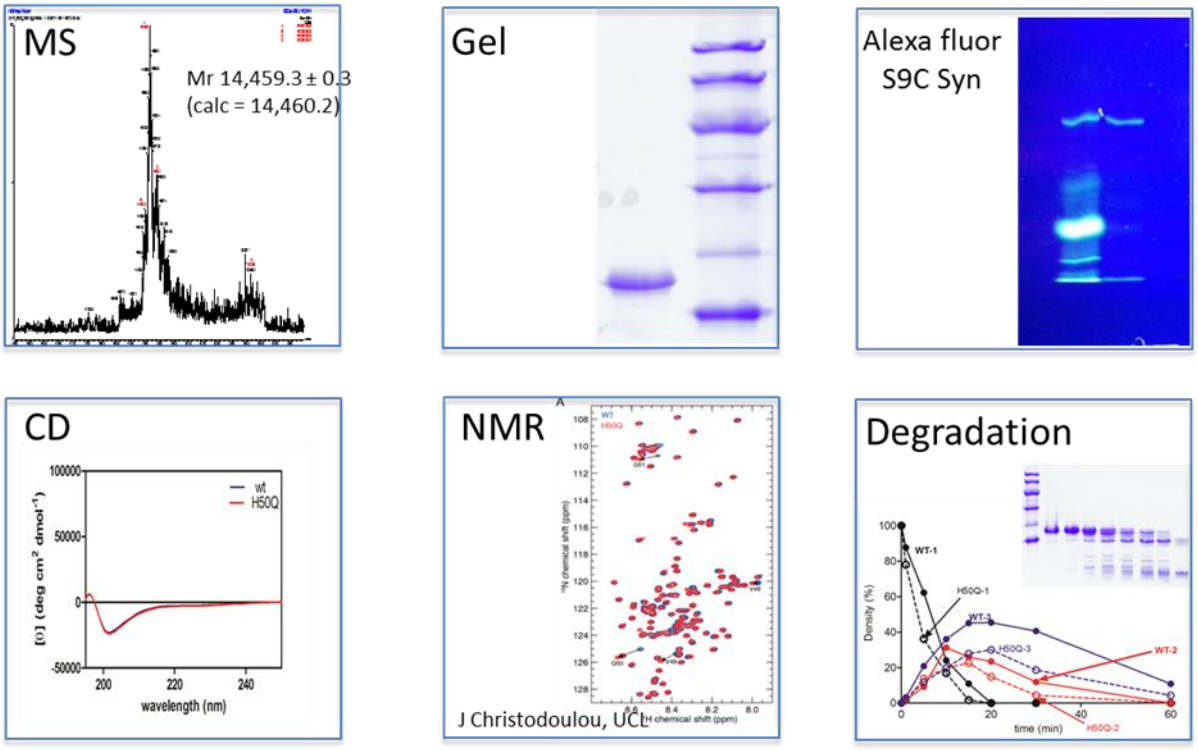


Figure 15. *Preliminary characterisation of α-synuclein to confirm sequence, chemical physical properties and stability*

After successful characterisation, a final large batch of α-synuclein and its N15 isotopically labelled protein was produced by UCL. Aliquots of the preparation were frozen after preliminary stability experiments and a pilot homogeneity study was carried out. An additional purification step was also carried out in small scale at LGC by reverse phase chromatography to reduce the amount of truncated forms of the protein present and to reach a purity above 98 % for α-synuclein.

Quantification traceable to the SI was performed by LGC by both amino acid analysis and using signature peptides which had been traceably quantified in-house. Amino acid analysis was performed by microwave assisted hydrolysis and double exact matching IDMS[28,29] with calibration against SI traceable amino acid standards sourced from NMIJ. The results from the amino acid analysis are reported in Figure 16.

Due to aggregation of α-synuclein during acid hydrolysis, the final quantification of the protein was carried out by tryptic digestion and using four peptides that were purity assessed and SI traceably quantified in house [30]. The selection of the peptides to be used for quantification was first carried out *in-silico.* This was followed by experiments to determine the theoretical ionisation efficiency, chromatographic properties, kinetic of release from the protein, and stability. The peptides selected for quantification were: MDVFMK (T1); QGVAEAAGK (T6); EGVLYVGSK (T8); TVEGAGSIAAATGFVK (T13). Quantification of the peptides was performed by using a purity corrected amino acid analysis approach[31]and confirmed by NMR. The results are reported in table 3.

Tryptic digest conditions were then optimised to obtain complete release of the peptides from the protein. The final protein concentration for α-synuclein was 4.3 ± 0.29 nmol/g.

Preliminary homogeneity studies were also carried out, and the solvent for diluting α-synuclein to the desired concentrations was optimised. The optimised solution was an aCSF consisting of Bovine Serum Albumin and electrolytes.

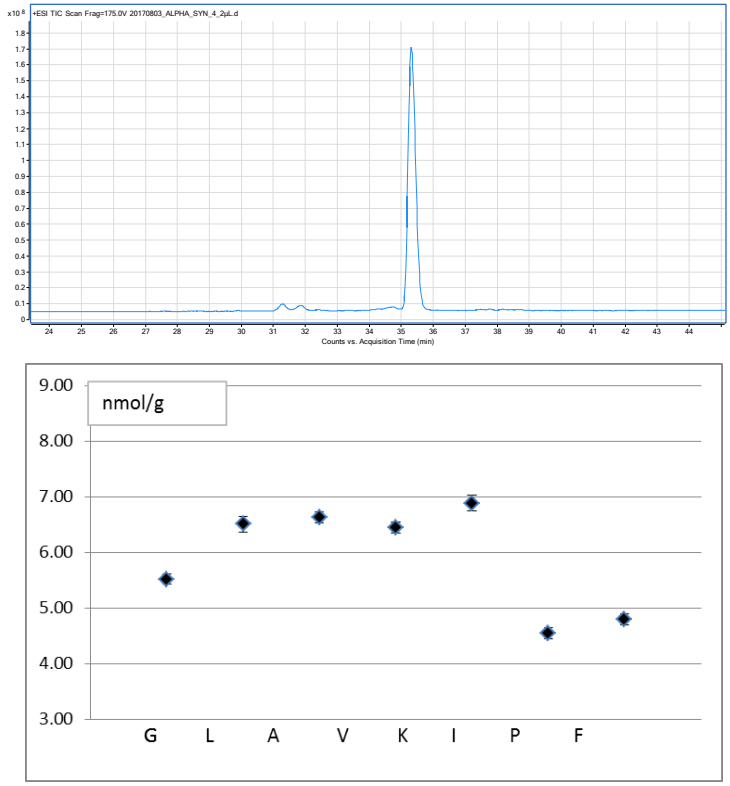


Figure 16. A. LC-MS profile of α-synuclein analysed as intact by LC-MS. B Content of glycine, leucine, alanine, valine, isoleucine, proline, and phenylalanine in the prepared α-synuclein solution. Error bars are the combined uncertainty of the amino acid analysis results (n=3).

	stated purity from manufacturer	Measured peptidic purity (%)	nmol/g of peptide solution (PICAA)
MDVFMK	≥95%	95.1 ± 0.6	21.01 ± 2.10
QGVAEAAGK	≥95%	78.7 ± 0.8	17.26 ± 4.28
EGVLYVGSK	≥95%	99.0 ± 0.2	28.92 ± 2.61
TVEGAGSIAAATGFVK	≥95%	26.1 ± 0.4*	4.96 ± 0.12

*Table 3. Results from the SI traceable quantification of the peptides to be used for quantification of the primary calibrator. The uncertainty reported is the expanded uncertainty (95 %). *prior additional reverse phase purification*

α-synuclein LC-MS method in CSF

A candidate RMP method was developed for quantification of α-synuclein in CSF based on tryptic digestion, high pH fractionation, and quantification by using a capillary LC-system coupled with a triple quadrupole mass spectrometer.

For method validation, quantification of (i) two equivalent pooled CSF aliquots purchased from SERA labs, prepared on two different days, and (ii) three aCSF + BSA samples spiked with α-synuclein at a concentration of 2 ng/g was performed using the candidate RMP against a single 5 point calibration curve. This gave an assessment of method repeatability, reproducibility and accuracy. This data, in addition to the primary calibrator quantification for α-synuclein, serve as an estimate of combined measurement uncertainty. After gravimetric correction, linearity of the quantification method was confirmed across the tested concentration range of 0 – 10 ng/g, which encompasses the expected endogenous concentration range of α-synuclein in CSF, see Figure 17.

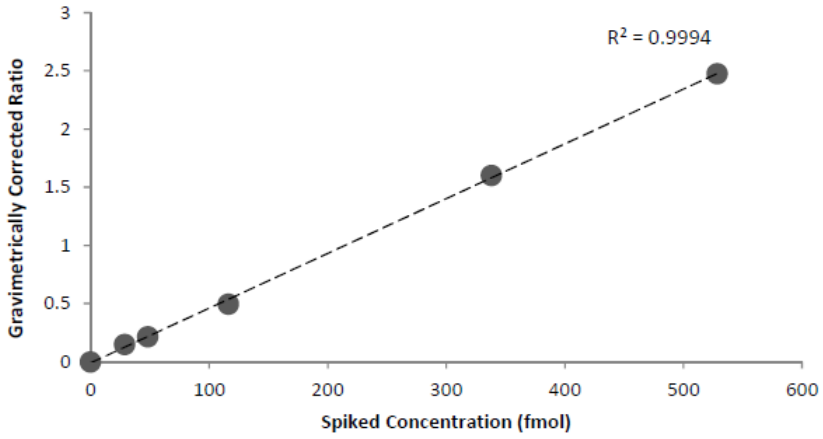


Figure 17. Example of calibration curve (T6).

Correlation curves for the peptides (T6, T8, T12 and T13) gave an $R^2 > 0.99$. The limit of quantification was at or below the lowest concentration calibration sample included (0.5 ng/g), with a relative standard deviation (RSD) of 6 %, 4 % and 15 % for analytical replicates for T6, 12 and T13 respectively.

Method repeatability and digestion reproducibility was found to be high, and a good agreement across the three peptides was observed with a triplicate quantification of a spiked aCSF sample giving an RSD of 2.3 % across all three samples and the three peptides quantified.

The CSF samples from cohort 1 and 10 CSF samples from PD patients prepared by CHU Mpt were also sent to LGC. These will be analysed in the follow-on project 18HLT09 NeuroMET2 to assess correlation of the candidate RMP with immunoassays, and the potential of the method to be used to assign SI traceable values to candidate reference materials.

α-synuclein uncertainty

The uncertainty of α-synuclein quantification in CSF by calibration curves was estimated by considering (i) the LC- Selected Reaction Monitoring quantitative analytical repeatability, (ii) digestion reproducibility and (iii) the uncertainty associated with the calibration curve itself. A combined total measurement uncertainty was then calculated by combining (i) the calibration curve uncertainty averaged across the three peptides, (ii) the uncertainty of the quantitative results across the three peptides and (iii) the uncertainty associated with the primary calibrator concentration.

The α-synuclein uncertainty from the calibration curve measurement was estimated based on the Eurachem guide to Quantifying Uncertainty in Analytical Measurement[32,33]. The standard deviation of the calibration curve residuals was calculated using Equation 3, where Aj is the i^{th} measurement, B_0 is the curve intercept, B_1 is the curve slope and c_j is gravimetric amount of α-synuclein added to the i^{th} calibration standard.

$$S = \sqrt{\frac{\sum_{j=1}^{n}[A_j - (B_0 + B_1 \cdot c_j)]^2}{n-2}}$$

Equation 3

The uncertainty associated with the calibration curve quantification was then estimated for each of the three samples and three peptides quantified using Equation 4, where C_0 is the measured α-synuclein sample concentration, $\bar{c}$ is the average concentration of all the calibration standards, p is defined as the number of analytical measurement performed on the sample and n is the number of calibration measurements.

$$u(c_0) = \frac{S}{B_1}\sqrt{\frac{1}{p} + \frac{1}{n} + \frac{(c_0 - \bar{c})^2}{S_{xx}}}$$

Equation 4

Using this approach, the project found that the two main contributors to the measurement uncertainty for α-synuclein were the calibration curve itself (9.2 %) and the primary calibrator stock concentration (6.8 %). This gave a final combined measurement uncertainty for α-synuclein of 12.0 %. It should be noted that this calculation does not include homogeneity of the sample.

Summary

Significant progress was made by the NeuroMET project towards the development of a total-tau (t-tau) and a α-synuclein reference measurement procedure (RMP) with traceability to the SI. However, validation of these RMPs could not be completed within the timeframe of this project and so will be continued in the follow-on project 18HLT09 NeuroMET2.

Recombinant t-tau 441 and the synthetic peptide GAAPPGQK, were SI traceably quantified using amino acid analysis, and their suitability as primary calibrators was evaluated for the quantification of t-tau in CSF. Increased variability was observed when using the synthetic peptide as an internal standard and therefore only recombinant t-tau 441 was selected for futher method development. A LC-MS method based on parallel reaction monitoring experiments was developed for the analysis of the peptides obtained after protein precipitation, solid phase extraction and tryptic digestion.

To evaluate the potential of the LC-MS method to be used as a RMP for standardisation of t-tau measurements as well as the commutability of potential tau certified reference materials with values assigned by using the RMP; three CSF pools with low, medium and high values of t- tau 441 were prepared and SI traceably quantified by using the NeuroMET RMP. The preliminary results are promising and the analysis of 40 additional samples will be carried out in the follow-on project 18HLT09 NeuroMET2.

An SI traceable LC-MS reference method was also developed for quantification of α- synuclein in CSF. Recombinant α-synuclein was prepared, purified, and SI traceably quantified by using three signature peptides and amino acid analysis. The primary calibrator was aliquoted and preliminary homogeneity studies were carried out. Results showed homogeneity (within 5 %) of the SI traceable quantified aliquots and when those aliquots were used for further fresh dilutions. A method for quantification of α- synuclein based on tryptic digestion, solid phase fractionation, and capillary-LC-MS experiments was also developed and validated. The limit of detection for this method was 0.5 ng/ml and the uncertainty was 12 %.

4.5 *Objective 5. To determine and characterise PCOM for NDD. This will be achieved by developing improved clinical assessment questionnaires focused on the decline in motor and cognitive functions and increases in behavioural, communicative and psychological symptoms (e.g. agitation) in AD patients. For PD patients, this will be also supported by the validation of a novel immunoassay for the measurement of the stress biomarker cortisol.*

A large number of PCOMs of cognitive performance (e.g. memory tests) are currently available for use in clinical practice and research settings, but their accuracy and sensitivity to discriminate clinical onsets has yet to be established. The translation of traditional metrological concepts into the field has the potential to be revolutionary, particularly in the assessment of memory decline which is the most prominent symptom in the early stages of the disease. Therefore, the aim of the project was to address this by developing and characterising PCOM for NDD.

RISE, in collaboration with ModusOutcomes, conducted several studies on the battery of neuropsychological assessments and self-reports of health outcomes (including legacy cognitive performance PCOMs) applied to cohort 1, on published data, and on data obtained from the Gothenburg MCI study. Based on this:

- Rasch Measurement Theory (RMT) was applied on several neuropsychological assessments for the first time
- a prototype NeuroMET Memory Metric was developed that will be validated in the follow-on project 18HLT09 NeuroMET2 for further implementation into clinics
- Construct Specification Equations (CSE) to be used as mathematical references were developed.

RMT separates estimates of person and item attribute values and their scaling on the same interval logit scale. The dichotomous case of the logistic regression function is shown in equation 5:

$$\log\left(\frac{P_{success}}{1-P_{success}}\right) = \theta - \delta$$

Equation 5

Rasch Measurement Theory (RMT)

In this project, RMT was considered to be a necessary pre-requisite for quality-assured measurements and the establishment of metrological traceability and primary references for categorical data in general.[34] For example, the RMT was applied to the Mini Mental State Examination (MMSE) data from cohort 1, and it was demonstrated that the data was not linear when using conventional approaches (i.e. such as the classical test theory (CTT) based on sum scores or t-scores; Figure 18). As a consequence, test persons with higher cognitive abilities were underestimated and test persons with lower cognitive abilities were overestimated. This limitation is particularly important for the MMSE test as it is frequently used in MCI patients. An example of the potentially serious impact of the distorted scale when correlating MMSE scores against brain volumes is illustrated in Figure 19[34,35].

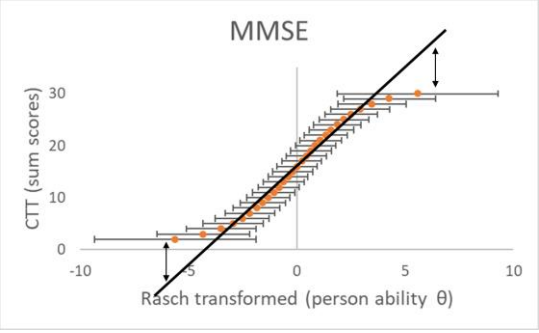


Figure 18. The classical test theory sum score on the y-axis and the Rasch-transformed person ability scores on the x-axis showing distortion towards the ends.

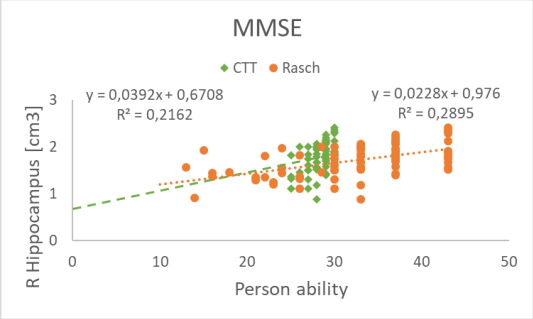


Figure 19 Correlation plot of right hippocampus volume (y-axis) and person ability (x-axis). The orange dots show Rasch-transformed values on the same scale as the CTT sum scores (green dots).

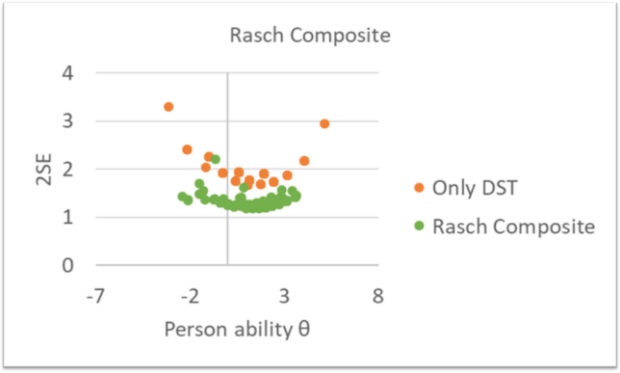


Figure 20. Measurement uncertainties for Digit Span Test (DST) in orange and a composite score including DST and other memory tests in green demonstrates. More and targeted items reduce the measurement uncertainties.

Figure 20 shows the advantages of using a Rasch composite model compared to individual tests, with the ultimate goal of reaching higher reliability by adding well targeted additional memory items. The largest gains in reliability with the composite model are found at the ends of the ability range, i.e. for persons with the lowest and highest abilities, where there are few items at the targeted difficulty levels. In particular, this has implications for accurate monitoring of memory ability in HC and MCI for early detection on memory decline.

Construct Specification Equations (CSE)

CSE were developed by RISE in collaboration with Modus Outcomes and used as reference to explain and improve memory tests[34,3619] CSEs offer three main advantages: i) an observational outcome (i.e. a response); ii) a causal mechanism that transmits variation in the intended attribute (e.g. memory test item difficulty), via an instrument to the observed response (e.g. the test person); and iii) attribute measures (i.e. the measurand) expressed in some unit. From a metrological point of view, CSEs resemble 'recipes for certified reference materials', familiar for providing traceability in chemistry and material properties. They also provide a comprehensive empirical understanding of: i) how a collection of items works together; ii) what is being measured; and iii) how validity is ensured.[Error! Bookmark not defined. Error! Bookmark not defined.] Moreover, CSEs provide a predictive tool for the design of, for example, new memory tasks complementing existing PCOMs and demonstrating item equivalence.

The concept of entropy can play a major role when formulating CSEs and in explaining and enabling improvement of memory tests.[34,35] Entropy can play important roles not only in explaining task difficulty but also person ability.[37] Therefore, RISE in collaboration with ModusOutcomes focused on the development of CSEs for the memory test item difficulty measures. Details of the CSE calculation and methodology measures developed within the project are in Melin et al. [2019] and Pendrill [2019]. [34,36] In summary, it includes five steps:

1. An RMT analysis of the data, which implies separate estimates of person and item attribute values and their scaling on the same interval logit scale. For the purpose of CSEs for task difficulty, the Rasch estimates, δ_j, for each item, j, were used
2. Identification of a set of explanatory variables, X_k, to inform a model of the expected item attribute
3. Principal component analysis (PCA) amongst the set of explanatory variables, X_k
4. Linear regression of the Rasch estimates, δ_j, [step 1] against X' in terms of the principal components, P [step 3], and the original data
5. Conversion back from principal components to the explanatory variables, X_k, in order to express the CSE for the item attribute in the form

Expanded measurement uncertainties were estimated ($k=2$), giving $U(\delta)$ through the RMT analysis [step 1], $U(X_k)$ and $U(\beta)$.

The initial evaluation of CSE was carried out on the Knox Cube Test (KCT), a PCOM that involves patients replicating different tapping series on four black one-inch cubes placed in a row. The set of items comprises different lengths of sequence, from 2 taps to 7 taps, and sequences with different orders. Figures 21a and 21b show an example where task difficulty (zR) for a sequence of increasing difficulty in the Knox cube test is explained as a sum of three terms: entropy, number of reversals and the average distance covered in each sequence. The specification equation:

$$zR_i = Intercept + 1(1) \cdot Entropy_i + 0.6(1.3) \cdot Reversals_i + 0.4(9)\ Distance_i$$

can be used to predict the difficulty of sequence *i*. In brackets are the measurement uncertainties with a coverage factor of 2. The same principles were applied to a number of neuropsychological assessments used to assess participants of cohort 1 and the results were CSEs similar to the one for KCT, both in terms of explanatory variables and predictability.

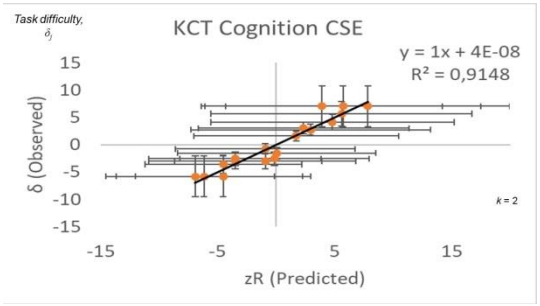


Figure 21a Linear regression of the measured task difficulty, *δ*, against the CSE estimates *zR* based on the three explanatory variables; *X = {Entropy, Reversals, Average Distance}* for the series of KCT sequences.

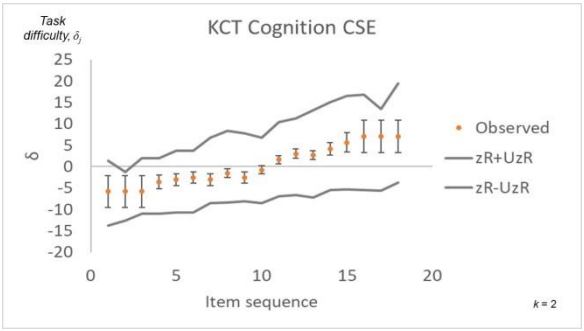


Figure 21b Predicted values of task difficulty, δ, for series of KCT sequences of increasing difficulty from a CSE, zR, based predominantly on entropy (with minor additional contributions from the number of reversals and average distance of each sequence, see Table 1) compared with corresponding measurement values (red dots with uncertainty intervals). The corridor of model uncertainties is shown as zR + UzR, uncertainty coverage factor k = 2.

For cohort 2, the cortisol measurements were assessed for each donor grouping (i.e. HC, AD and PD) irrespective of the timepoint within each day, as well as for each timepoint between days. Using this simplified analysis, it was apparent that the coefficient of variation (CV) for the cortisol measurements were generally reduced when the percentage of free cortisol was assessed if compared with the CVs of the data grouped in the same manner for each donor. This suggests that the percentage of free cortisol normalises some of the variability associated with the time and day effects. When assessing the total cortisol measurements, it was difficult to identify any trends. However, when the percentage of cortisol in serum was derived and plotted, it was clear that the PD cohort exhibited a propensity for a higher level of cortisol than the other cohorts (see Figure 22), albeit with a small number of outliers for the control cohort (C). These were ultimately attributed to medications administered for treatment of hypertension and high cholesterol.

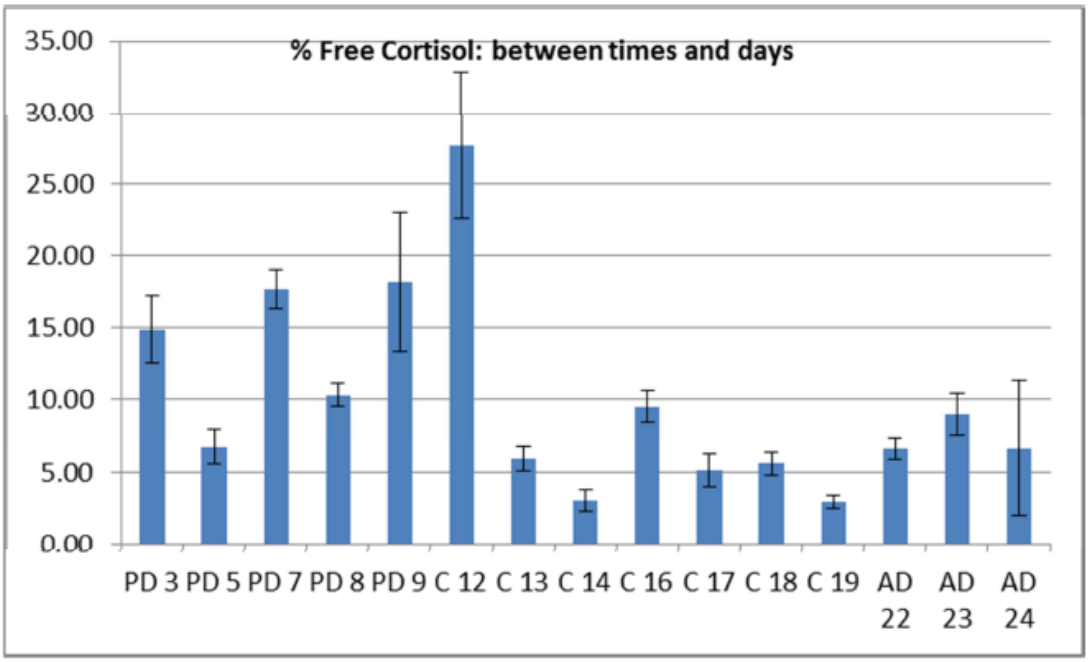


Figure 22. Mean value of the percentage of free cortisol measurements for each donor irrespective of the timepoint (11am, 1pm and 3 pm) for sample collection and the day of collection (days 1, 2 and 3).

A mixed effects model with maximum likelihood estimation was fitted to the responses from the HADS questionnaire from cohort 2 and the logarithm of the total and free cortisol measurements, using donor type, day, time, anxiety and depression as the fixed effects, and the patient and measurement variability as the random effects. No significant difference was found, principally due to the high patient-to-patient variability and the small number of samples. PD patients however showed a higher proportion of free cortisol than the control and AD groups; although this was not significant at the 95 % level (t = 1.62).

It is worth noting that this study has revealed that there was no time effect associated with the total cortisol measurements for all donors, and that there was neither time effect nor day effect associated with the percentage of free cortisol for all donors. This latter point supports the working hypothesis that the percentage of free cortisol normalises the variability in the absolute values of the free or total cortisol measurements between distinct days. In addition, within the time frame assessed (i.e. 11 am to 3 pm), there was no time effect on the measurements of total cortisol, only a day effect for the measurements for all donors. This observation may not be unexpected given that the greatest changes in the circadian rhythm occur much earlier in the morning than when the project's sample collection commenced; i.e. earlier than when routine sample collections may take place in clinics. However, the development of a point of care (POC) device would be more desirable for this study as the cortisol levels within the donors could be monitored without the need of a research nurse *in situ* to collect all the samples. Hence the scope of developing a magnetic bead-based solution phase assay was assessed. The ability to measure total cortisol in serum using a solution phase assay was successfully demonstrated, and this may be amenable for integration into a POC device in future studies.

Summary

Currently, there exist many PCOMs of cognitive performance (e.g. memory tests), which are commonly used both in clinical practice and clinical research. But these can neither claim accuracy and sensitivity to distinguish between patients (especially in early stage disease) nor are they metrologically proven. In this project, several studies were carried out on (i) the extensive battery of legacy cognitive performance PCOMs applied to the participants in cohort 1, (ii) published data, and (iii) data obtained from the Gothenburg MCI study. The process of quality assured measurement of cognitive performance was performed through: (i) applying the Rasch Model Theory on collective prospective patient PCOM data; (ii) development of a prototype NeuroMET Memory Metric based on legacy cognitive PCOMs; (iii) formulation of construct specification equations (CSEs) to link cognitive assessment outcomes to metrological concepts.

Studies were performed on the correlation of the NeuroMET data from cohort 1 with cognitive task difficulty and instrument parameters (e.g. sequence entropy) and it was shown that the mathematical models developed in ths project and based on the Rasch Model Theory can be used to predict for example task difficulty. Figures 20a and 20b show an example where task difficulty (zR) for a sequence of increasing difficulty in the Knox cube test is explained as a sum of three terms: entropy; number of reversals and the average distance covered in each sequence. The specification equation (below) can be used to predict the difficulty of sequence *i*.

$$zR_i = \text{Intercept} + 1(1) \cdot \text{Entropy}_i + 0.6(1.3) \cdot \text{Reversals}_i + 0.4(9)\ \text{Distance}_i$$

In brackets are the measurement uncertainties with a coverage factor of 2.

Free and total cortisol as a stress marker were also measured in serum samples from cohort 2. The results were promising, and it was possible to discriminate between PD and non-PD (i.e. AD and HC patients combined). However, limited overall conclusions could be drawn due to the small size of the patient cohort.

4.6 *Objective 6. To develop, validate and verify multimodal statistical analyses used to correlate NDD patient health status with AD biomarker and MRI data. These novel analyses will identify the most promising tools for NDD early diagnosis and disease progression monitoring, aiming for better resolution in detecting the early signs of disease and reducing the number of biomarker studies required (e.g. in cases where biomarker studies are especially challenging).*

In this project, we have extended the work of CSE focusing not only on task difficulty as previously, but also attempting to explain person cognitive ability, where the CSE could be a function of diverse set of volumes of AD-related brain structures and neurometabolite concentrations. Thus, a CSE for patient cognitive ability would be able to provide better metrological assurance of biomarker analyses compared to univariate studies.

Faced with the relatively large measurement uncertainties of not only the PCOMs but also each biomarker, a critical figure of merit metrologically is an indication of the responsiveness of the various explanatory variables *X* in the CSE which can be truly detected. The impact of assessing smallest detectable change (SDC) is to provide the end-user with estimates of the current responsiveness of the tools to date for measuring both cognitive task difficulty and person ability.

RISE, in collaboration with Modus Outcomes, developed for the first time CSEs where person cognitive (memory) ability was described as a function of several explanatory variables such as biomarkers as well as a model for calculating SDC in those CSEs.

Person ability Construct Specification Equations (CSE)

Formulation of CSEs for person ability, θ, follows a substantially similar process as described above and exemplified for task difficulty, δ. Faced with a plethora of biomarkers which can be correlated with a person's cognitive ability, a choice of the most promising biomarkers was made on the basis of the following criteria:

- Degree of correlation amongst biomarkers using PCA according to the procedure described for CSE for task difficulty. A cluster of statistically correlated biomarkers (typically with Pearson coefficients above 0,4) may indicate that it would be sufficient to choose just one of the biomarkers in each such cluster
- Initial univariate correlation plots of each person ability, θ, against a selection of biomarkers. A statistically significant degree of correlation (typically with Pearson coefficients above 0,4) for a particular biomarker (preferably without correlation with other biomarkers) might indicate a promising candidate
- A full multivariate PCR according to the procedure described for CSE for task difficulty

A model for calculating smallest detectable change (SDC)

The SDCs are limited in the first case by the measurement uncertainties in the psychometric attribute values $u(\delta)$ derived with the RMT eq. (5). A simple derivation of these 'distribution-based' metrics yields the following expression for the SDC associated with the explanatory variable X [Melin et al 2019a]:

$$SDC(\beta_X \cdot X) = 2 \cdot \sqrt{2} \cdot u(\delta) \Rightarrow SDC(X) = \frac{2 \cdot \sqrt{2} \cdot \overline{u(\delta)}}{\beta_X}$$

Equation 6

The quoted uncertainties (coverage factor *k* = 2) in each SDC have been calculated from the expression for the standard uncertainty:

$$u[SDC(X)] = \frac{2 \cdot \sqrt{2} \cdot \overline{u(\theta)} \cdot u(\beta_X)}{(\beta_X)^2}$$

where $u(\beta_X)$ is the modelling uncertainty $\mathbf{u}(\boldsymbol{\beta}) = \mathbf{P} \cdot \mathbf{u}(\hat{\mathbf{C}})$, provided by the MathCad built-in function *polyfit,* which was used to perform the regression analysis.

The responsiveness of each CSE can be judged by evaluating equation (6) for the sets of explanatory variables. For example, for the KCT item sequence 2-1-4, entropy corresponds to ~1.8 and the SDC for entropy is 1.5(5) This implies that the SDC is smaller than the explanatory variable, indicating that this variable is truly detected rather than 'buried in the noise' of measurement uncertainty. At this stage, corresponding SDCs for

person abilities were found to be greater than the value of their associated explanatory variables, indicating that the explanatory variable can barely be detected, i.e. the instrument is not sufficiently responsive.

Figure 23 shows the propagation of measurement uncertainties, portrayed as an Ishikawa diagram, from the initial neuropsychological assessment, through Rasch analysis; formulation of CSE for both cognitive person ability (upper half) and task difficulty (lower half); and finally, estimation of SDC for each explanatory variable and Rasch attribute.

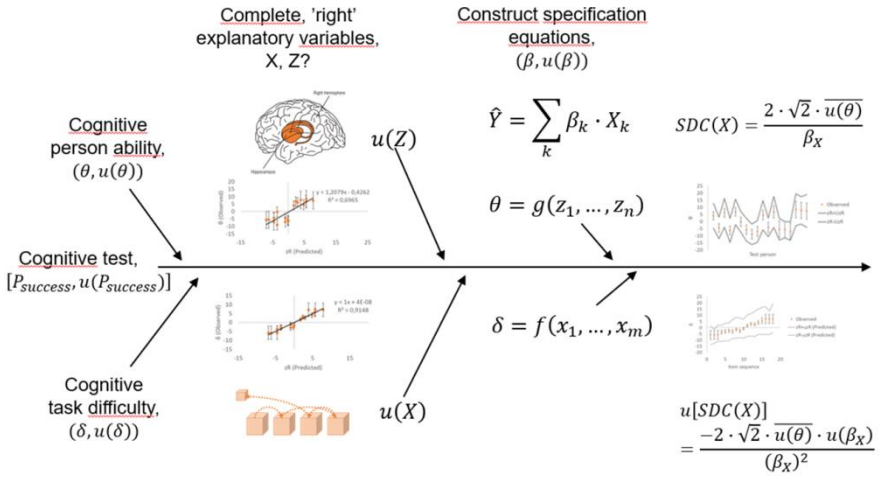


Figure 23 Propagation of measurement uncertainties, portrayed as an Ishikawa diagram, from the initial neuropsychological assessment, through RMT analysis; formulation of CSE for both cognitive person ability (upper half) and task difficulty (lower half); and finally, estimation of SDC for each explanatory variable and Rasch attribute.

Summary

The project developed for the first time a set of CSE in the field of cognitive assessment and NDD where person cognitive (memory) ability is described as a function of several explanatory variables such as biomarkers. The CSE enable improved understanding and prediction of both cognitive task difficulty and person ability. Furthermore, their potential to be used as reference in the field will for the first time introduce metrological and tracebility concepts into PCOM evaluation. The longitudinal studies in the follow-on project 18HLT09 NeuroMET2 will further consolidate the applicability of CSE for standardisation of PCOM.

Multivariate principal component regression was also applied to the NeuroMET data from cohort 1 for all the biomarkers from objective 3 in plasma, CSF and saliva (Aβ40, Aβ42, t-tau, NFL, α-synuclein and cortisol) together with the MRI/MRS data from objective 2. Formulation of causal Rasch models as prototype metrological references for cognition was initially guided by data from other studies relating to disease state indices, while using the NeuroMET developed cognitive assessment instruments (specially for memory).

The results of this have shown the potential of the procedure to define a metrological uncertainty for the diagnosis of NDD patients.

5 Impact

In this project a multidisciplinary metrological infrastructure was for the first time developed to provide RMP to improve accuracy in diagnosis of NDD and particularly AD by addressing the key steps involved in AD disgnosis including cognitive assessment, MR and fluid biomarkers. Thus, providing a pool of data that together with clinicial expertise will help to define patient diagnosis and outcomes.

The NeuroMET partners attended a large number of technical and NDD diagnostic conferences over the course of the project and contributed with posters and oral presentations. A summary of the final results from the project were presented at key conferences e.g. the Alzheimer Association International Conference (2019), the IFCC EUROMEDLAB Conference (2019), and IMEKO (2019) targeting the clinical, standardisation and metrological communities. See the project website for more details https://www.lgcgroup.com/our-programmes/empir-neuromet/

The consortium also organised three stakeholder meetings in Sweden, Germany and France that were attended by academic and clinical stakeholders and targeted all aspects of the project. All stakeholder meetings we formed by three sections on cognitive assessment, MR and fluid biomarkers and were followed by a discussion with stakeholders.

In addition, project partners attended and organised a number of workshops and training course addressing quality assurance in person centred healthcare and cognitive assessments for industry and broader scientific community, PCOM (objective 5), Rasch analysis for clinicians (objective 6) and MRS data analysis also for clinicians (objective 2).

Activities from objective 2 (7T MRI and MRS) and objective 4 (t-tau) have also been positively integrated with NDD standardisation activities e.g. EUFIND (European Ultrahigh-Field Imaging Network for Neurodegenerative Diseases) and The International Federation of Clinical Chemistry and Laboratory Medicine Working Group on CSF-proteins (IFCC WG CSF) Standardisation.

Impact on industrial and other user communities

Metrological quality assurance of PCOMs is essential to underpin reliable clinical decisions for AD diagnosis and recruitment in clinical trials. Within NeuroMET an increased precision when measuring patient memory ability was demonstrated when processing data by applying the Rasch Model (from objective 6) in comparison with traditional methods. Two training courses were organised at partner Charité in Berlin and at Kristianstad University (PMHealth). The first training course was to initiate the familiarisation of clinicians with the project's Rasch Model Theory and the second training course addressed the broader scientific community. The methodologies used to develop improved cognitive assessment protocols (objective 5) were also presented to clinical end users via seminars, conferences and a number of peer review publications.

In addition, a patient group meeting targeting cohort 2 was organised with the aim of informing patients of the results of the project and the progress made in the field. The meeting attended by the lay advisory panel for cohort 2 included AD and PD patients. The participants expressed much interest in the results from the project and potential participation in follow-on studies.

The improved cognitive clinical tests and rating scales (objective 5) developed within the NeuroMET project will provide guidance and tools to clinicians for improved diagnosis and better prediction of future NDD patient decline. The prototype Memory Metric (objective 5) will be validated in the follow-on project 18HLT09 NeuroMET2 through longitudinal studies. Three training courses were organised addressing predominantly industry stakeholders. They were on (1) Assuring quality in person-centred healthcare; (2) How to calibrate a questionnaire: quality assuring categorical data with psychometric measurement theory and (3) Quality assured categorical data.

The *in-vitro* diagnostic (IVD) industry will also benefit from the results of this project by gaining additional information on instrument performance (objective 3) and by the availability of the developed reference measurement procedures (objective 4), which will facilitate regulatory approaval of new instruments and compliance with the IVD Directive 98/79/EC. Transfer into a clinical laboratory of the SI traceable LC-MS reference method for quantification of α-synuclein (objective 4) is currently on-going.

Impact on the metrological and scientific communities

The NeuroMET project has for the first time established important links between the metrological community and the NDD community through engagement in international NDD standardisation initiatives:

- round robin studies were organised with the IFCC WG CSF to underpin standardisation of t-tau with the first comparison starting in September 2019. In the comparison the NeuroMET primary calibrator for t-tau (from objective 4) will be shared between all participants to support standardisation of the results.
- within the EUFIND consortium NeuroMET partners were involved in the harmonisation of MRS protocols for application in NDD research cohorts and investigation of comparability of MRS data acquired at different 7T research sites and across different vendor platforms (using findings from objective 2).

The project's progress and approaches developed for the purity assessment of primary calibrators (objective 4) were regularly discussed within the BIPM Consultative Committee for Amount of Substance: Metrology in Chemistry and Biology (CCQM) metrological community. Indeed, the methods developed through NeuroMET to characterise the α-synuclein primary calibrator (also objective 4) have enabled the first ever Calibration Measurement Capability claim, for partner LGC, on purity determination of peptides.

Impact on relevant standards

The NeuroMET consortium partners were involved in and presented the project results to ISO technical committees such as TC 12 on Quantities and Units, TC 212 on Clinical laboratory testing and in-vitro diagnostic test systems, TC 215 on Health Informatics. A new ISO TC 276 Biotechnology WG3 Analytical Methods project aimed at identifying protein analytical standardisation requirements for advanced therapeutics and biotechnology, has also been initiated, and this will incorporate selected NeuroMET methods for peptide purity (from objective 4) and for structural analysis. The project was also presented at the JCTLM (Joint Committee for Traceability in Laboratory Medicine) workshop on Accurate Results for Patient Care.

Longer-term economic, social and environmental impacts

The NeuroMET project successfully combined the strengths of NMIs and NDD clinicians to establish a metrology infrastructure that will provide measurement guidance and reference tools to NDD clinicians, academics, and pharmaceutical companies for:

- Appropriate study design and definition of the uncertainty of NDD clinical assessment protocols to improve diagnosis and NDD progression monitoring;
- More accurate stratification of patient cohorts, with respect to NDD status, for enrolment into clinical trials and for more informed patient management;
- Improved measurement comparability for NDD biomarkers through optimised measurement procedures and development of SI traceable reference methods for key biomarkers.

Going forward, the follow-on project 18HLT09 NeuroMET2 will positively contribute to the uptake and validation of this project's outputs through longitudinal studies, as well as further engagement with instrument manufactures and international activities such as through the JCTLM, IFCC and EUFIND.

6 List of publications

[1]. S. Pang and S. Cowen, 2018, Scientific Reports, https://doi.org/10.1038/s41598-017-17823-y

[2]. G. Franceschi, 2017, JBC http://www.jbc.org/content/292/17/6927.full.pdf

[3]. Cano S, Pendrill L, Barbic S, Fisher WP, 2017, Journal of Physics, Conference Series, https://doi.org/10.1088/1742-6596/1044/1/012057

[4]. Cano S, Pendrill L, Barbic S, Fisher WP, 2017, Journal of Physics, Conference Series, https://doi.org/10.1088/1742-6596/1065/7/072033

References

[1] Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67*(6), 361-370.

[2] Forjaz, M. J., Rodriguez-Blázquez, C., Martinez-Martin, P. et al , (2009), Rasch analysis of the hospital anxiety and depression scale in Parkinson's disease . *Mov. Disord*., 24: 526-532. doi:10.1002/mds.22409

[3] Leentjens, A. F., Dujardin, K., Pontone, G. M., Starkstein, S. E., Weintraub, D. and Martinez-Martin, P. (2014), The Parkinson Anxiety Scale (PAS): Development and validation of a new anxiety scale. *Mov Disord*., 29: 1035-1043. doi:10.1002/mds.25919

[4] SHANKAR K.K., WALKER M.D., FROST & M. W. ORRELL (1999) The development of a valid and reliable scale for rating anxiety in dementia (RAID), *Aging & Mental Health*, 3:1, 39-49, DOI: 10.1080/13607869956424

[5] Marques J. P., Kober T., Krueger G., van der Zwaag W., Van de Mortele P.-F., Gruetter R. R. (2010). 'MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field', *NeuroImage* 49:1271-1281

[6] O'Brien K. R., Kober T., Hagmann P., Maeder P., Marques J., Lazeyras F., Krueger G., Roche A.(2014)Robust T1-Weighted Structural Brain Imaging and Morphometry at 7T Using MP2RAGE'. *PLOS One* 9:e99676

[7] Gaser C., Dahnke R.. (2016) 'CAT – A Computational Anatomy Toolbox for the Analysis of Structural MRI Data.', *HBM* 2016:33-348

[8] Fischl B.(2012) 'FreeSurfer', *NeuroImage* 62:774-781

[9] Raz N., Lindenberger U., Rodrigue K. M., Kennedy K. M., Head D., Williamson A., Dahlke C., Gerstorf D., D. Acker J. (2005). 'Regional Brain Changes in Aging Healthy Adults: Genderal Trends, Individual Differences and Modifiers', *Cerebral Cortex* 15:1676-1689.

[10] Voevodskays O., Simmons A., Nordenskjöld R., Kullberg J., Ahlström H., Lind L., Wahlund L.-O., Larsson E.-M., Westman E., and ADNI (2014). 'The effects of intracranial volume adjustment approaches on multiple regional MRI volumes in healthy aging and Alzheimer's disease'. *Front Aging Neurosci* 6:264

[11] Iacono M. I., Neufeld E., Akinnagbe E., Bower K., Wolf J., Vogiatzis Oikonomidis I., Sharma D., Lloyd B., Wilm B. J., Wyss M., Pruessmann K. P., Jakab A., Makris N., Cohen E. D., Kuster N., Kainz W., Angelone L. M. (2015) 'MIDA: A Multimodal Imaging-Based Detailed Anatomical Model of the Human Head and Neck', *PLoS ONE* 10:e0124126

[12] Mlynarik V., Gambarota G., Frenkel H., Gruetter R. (2006). 'Localized Short-Echo-Time Proton MR Spectroscopy With Full Signal-Intensity Acquisition'. *Magn Reson Med* 56:956-970

[13] Mekle R., Mlynarik V., Gambarota G., Hergt M., Krueger G., Gruetter R. R. (2009). 'MR Spectroscopy of the Human Brain With Enhanced Signal Intensity at Ultrashort Echo Times on a Clinical Platform at 3T and 7T'. *Magn Reson Med* 61:1279-1285

[14] Wang H, Tan L, Wang HF, Liu Y, Yin RH, Wang WY, Chang XL, Jiang T, Yu JT. (2015) 'Magnetic Resonance Spectroscopy in Alzheimer's Disease: Systematic Review and Meta-Analysis'. *J Alzheimers Dis* 46:1049–1070. doi: 10.3233/JAD-143225

[15] Riese F, Gietl A, Zölch N, et al (2015). 'Posterior cingulate GABA and glutamate+glutamine are reduced in amnestic mild cognitive impairment and are unrelated to amyloid deposition and APOE genotype'. *Neurobiol Aging* 36:53–59. doi: 10.1016/j.neurobiolaging.2014.07.030

[16] Bell S. (2001) 'Measurement Good Practice Guide No. 11 (Issue 2) - A Beginner's Guide to Uncertainty of Measurement', National Physical Laboratory, ISSN 1368-6550

[17] Moeller S, Yacoub E, Olman CA, Auerbach E, Strupp J, Harel N, Ugurbil K (2010) 'Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI'. *Magn Reson Med* 63: 1144-53

[18] Pang S. & Cowen S. (2017) 'A generic standard additions based method to determine endogenous analyte concentrations by immunoassays to overcome complex biological matrix interference', *Scientific Reports*

[19] Vanderstichele H et al. (2012) Alzheimers Dement. Jan;8(1):65-73.

[20] Le Bastard N. et al. (2015) Clin Chemistry, 2015, 61:5, 734

[21] Polymeropoulos, M.H., Lavedan, C., Leroy, E., Ide, S.E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R. and Stenroos, E.S. (1997). Mutation in the α-synuclein gene identified in families with Parkinson's disease. *science*, *276*(5321), 2045-2047.

[22] Spillantini, M.G., Schmidt, M.L., Lee, V.M.Y., Trojanowski, J.Q., Jakes, R. and Goedert, M. (1997). α-Synuclein in Lewy bodies. *Nature*, *388*(6645), 839.

[23]Bousset, L., Pieri, L., Ruiz-Arlandis, G., Gath, J., Jensen, P.H., Habenstein, B., Madiona, K., Olieric, V., Böckmann, A., Meier, B.H. and Melki, R. (2013). Structural and functional characterization of two alpha-synuclein strains. *Nature communications*, *4*, 2575.

[24] Bergström, J. and Ingelsson, M. (2016). Alpha-Synuclein as a Diagnostic Biomarker for Parkinson's Disease. In *Immunotherapy and Biomarkers in Neurodegenerative Disorders* 215-234. Humana Press, New York, NY.

[25] Olanow, C.W. and Kordower, J.H. (2017). Targeting α-Synuclein as a therapy for Parkinson's disease: The battle begins. *Movement Disorders*, *32*(2), 203-207.

[26] Mollenhauer, B. (2014). Quantification of α-synuclein in cerebrospinal fluid: How ideal is this biomarker for Parkinson's disease?. *Parkinsonism & related disorders*, *20*, S76-S79.

27 Mollenhauer, B., Bowman, F.D., Drake, D., Duong, J., Blennow, K., El-Agnaf, O., Shaw, L.M., Masucci, J., Taylor, P., Umek, R.M. and Dunty, J.M. (2019). Antibody-based methods for the measurement of α-synuclein concentration in human cerebrospinal fluid–method comparison and round robin study. *Journal of neurochemistry*, *149*(1), 126-138.

[28] Cryar, A., Pritchard, C., Burkitt, W., Walker, M., O'Connor, G., Burns, D.T. and Quaglia, M. (2013). Towards absolute quantification of allergenic proteins in food—lysozyme in wine as a model system for metrologically traceable mass spectrometric methods and certified reference materials. *Journal of AOAC International*, *96*(6), 1350-1361.
[29]Torma, A.F., Groves, K., Biesenbruch, S., Mussell, C., Reid, A., Ellison, S., Cramer, R. and Quaglia, M. (2017). A candidate liquid chromatography mass spectrometry reference method for the quantification of the cardiac marker 1-32 B-type natriuretic peptide. *Clinical Chemistry and Laboratory Medicine (CCLM)*, *55*(9), 1397-1406
[30] Pritchard, C., Groves, K.J., Biesenbruch, S., O'Connor, G., Ashcroft, A.E., Arsene, C., Schulze, D. and Quaglia, M. (2014). Quantification of human growth hormone in serum with a labeled protein as an internal standard: essential considerations. *Analytical chemistry*, *86*(13), 6525-6532.
[31] Josephs, R.D., Stoppacher, N., Westwood, S., Wielgosz, R.I., Li, M., Quaglia, M., Melanson, J., Martos, G., Prevoo, D., Wu, L. and Scapin, S. (2017). Concept paper on SI value assignment of purity-Model for the classification of peptide/protein purity determinations. *Journal of Chemical Metrology*, *11*(1).
[32] https://www.eurachem.org/images/stories/Guides/pdf/QUAM2012_P1.pdf
[33] https://www.bipm.org/utils/common/documents/jcgm/JCGM_100_2008_E.pdf
[34] Pendrill L. R., (2019) Quality Assured Measurement - Unification across Social and Physical Sciences, *Springer Series in Measurement Science and Technology*, DOI: 10.1007/978-3-030-28695-8
[35] Pendrill L R, 'Assuring measurement quality in person-centred healthcare', *Measurement Science & Technology,* Volume 29, Number 3 *, 034003 special issue Metrologie 2017*, 2018, https://doi.org/10.1088/1361-6501/aa9cd2
[36] Melin J, Cano S, Pendrill L, (2019), Perspectives Sci., Submitted
[37] Pendrill L. R., Melin J., Cano S and the NeuroMET consortium 2019, 'Metrological references for health care based on entropy', *19th International Congress of Metrology, Paris (FR)*, EDP Science: web of conference open access, https://cfmetrologie.edpsciences.org/component/issues/, in press