MES-CoBraD

Multidisciplinary Expert System
for the Assessment & Management
of Complex Brain Disorders

D3.1

Project Manual – CoBraD RWD, Updated MES-CoBraD Protocols and Quality and Quantity (Q&Q) Evaluation – v1

June 2021

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

www.mes-cobrad.eu

PREFACE

The Multidisciplinary Expert System for the Assessment & Management of Complex Brain Disorders(MES-CoBraD) is an interdisciplinary project combining Real-World Data (RWD) from multiple clinical and consumer sources through comprehensive, cost-efficient, and fast protocols towards improving diagnostic accuracy and therapeutic outcomes in people with Complex Brain Disorders (CoBraD), as reflected in Neurocognitive (Dementia), Sleep, and Seizure (Epilepsy) disorders and their interdependence.

1. It brings together internationally recognized experts in medicine, engineering, computer science, social health science, law, and marketing and communication from across Europe, and combines clinical information and scientific research in CoBraD with technical innovation in secure data-sharing platforms, artificial intelligence algorithms, and expert systems of precision and personalized care, with a primary focus on improving the quality of life of patients, their caregivers, and the society at large.

2. It leverages RWD from diverse CoBraD populations across cultural, socioeconomic, educational, and health system backgrounds, with special attention on including vulnerable populations and minorities in an equitable manner and engaging key stakeholders to maximize project impact.

3. The project will deliver a rigorous and self-standing methodology that will drive the MES-CoBraD implementation and define its operational principles; The MES-CoBraD solution, through the implementation of novel, self-standing AI based components and their integration under a common platform for scientific exploitation will assist focusing on the evaluation and validation of the solution, the spread of the excellence gained, the expansion of its ecosystem and the real-life sustainability.

CONSORTIUM

NTUA	National Technical University of Athens	EL
NIA	Neurological Institute of Athens	EL
HSCSP	Fundació privada Institut de Recerca de l'Hospital de la Santa Creu i Sant Pau	IR
VUB - LSTS	VRIJE UNIVERSITEIT BRUSSEL	BE
ups	UPPSALA UNIVERSITY	SE
RMC-C&E	Clalit Health Services- Rabin Medical Center Cognitive Neurology and Epilepsy Clinics	IL
kcl	King's College London	UK
HOLISTIC	HOLISTIC P.C.	EL
SIMAVI	Software Imagination & Vision	RO
LIBER	STICHTING LIBER	NL
ENG	Engineering - Ingegneria Informatica S.p.A.	EN
EVOLUTION	MICHOPOULOS I. & CH. G.P.	EL
UoE	University of Edinburgh	UK
CEL	CyberEthics Lab	IT

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

MULTIDISCIPLINARY EXPERT SYSTEM FOR THE ASSESSMENT & MANAGEMENT OF COMPLEX BRAIN DISORDERS

GA#:	**965422**	**Start Date:**	**01/04/2021**
Topic:	**SC1-DTH-12-2020**	**Duration:**	**36 months**
Type of Action:	**RIA**	**Coordinator:**	**NTUA**

Deliverable Number	3.1
Deliverable Title	**Project Manual – CoBraD RWD, Updated MES-CoBraD Protocols and Quality and Quantity (Q&Q) Evaluation – v1**
Work Package Number	WP 3
Task Number	T3.2-T3.4
Date of Delivery	30/06/2021
Dissemination Level	Public
Work Package Leader	RMC
Lead Beneficiary	RMC
Authors	RMC
Contributors	KCL, NIA, SP, UU, ENG
Reviewers	ENG, NIA

Disclaimer

Copyright Message

Document History

Version	Date	Author (Partner)	Remarks/Changes
0.10	25/4/2021	Ophir Keret (RMC)	ToC
0.20	08/06/2021	Ophir Keret (RMC) Elissaios Karageorgeou (NIA) Victor Montal (SP) Luiz Eduardo Mateus Brandao (UU) Antonio Valentin Huete (KCL) Ioannis Stavropoulos (KCL) Antonino Sirchia (ENG)	First Draft
0.23	22/06/2021	Ophir Keret (RMC) Elissaios Karageorgeou (NIA) Victor Montal (SP) Luiz Eduardo Mateus Brandao (UU) Antonio Valentin Huete (KCL) Ioannis Stavropoulos (KCL) Antonino Sirchia (ENG)	Second draft
0.25	24/06/2021	Ophir Keret (RMC) Elissaios Karageorgeou (NIA) Victor Montal (SP) Luiz Eduardo Mateus Brandao (UU) Antonio Valentin Huete (KCL) Ioannis Stavropoulos (KCL) Antonino Sirchia (ENG)	Draft ready for review
0.30	28/06/2021	Elissaios Karageorgeou (NIA)	Review 1
0.40	28/06/2021	Antonino Sirchia (ENG)	Review 2
0.50	28/06/2021	Ophir Keret (RMC)	Consolidation of changes from the reviews
0.60	29/06/2021	Ophir Keret (RMC)	Final edit
0.70	30/06/2021	Konstantinos Alexakis (NTUA) Michael Kontoulis (NTUA)	Quality Control
1.00	30/06/2021	Christon Ntanos (NTUA)	Final Version to be Submitted

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Table of Contents

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

LIST OF FIGURES

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

LIST OF TABLES

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

LIST OF ABBREVIATIONS

Abbreviation / Acronym	Description
µl	*Microliter*
AB	*Advisory Board*
API	*Application programming interface*
BOLD	*Blood Oxygen Level Dependent*
CC	*Chief Complaint*
CCC	*Chronic Complex Condition*
CEL	*CyberEthics Lab*
CERAD	Consortium to Establish a Registry for Alzheimer's Disease
cm	*Centimeter*
CoBraD	*Complex Brain Disorders*
CSD	*Complex Structure Real World Data*
CSH	*Clinical Symptoms and History*
DtP	*Direct to Platform*
Dx.y	*Deliverable number y belonging to WP x*
EB	*Ethics Board*
EC	*European Commission*
ECLIA	*Electro Chemi Luminescence Immuno Assay*
EDF	*European Data format*
EDTA	*Ethylenediaminetetraacetic acid*
EEG	*Electroencephalograpy*
EMR	*Electronic Medical Records*
EU-GDRP	*European General Data Protection Regulation*
FH	*Family History*
FLAIR	*FLuid Attenuated Inversion Recovery*
FS	*Functional status*
HPA	*Hypothalamus-Pituitary-Adrenal*
JPEG	*Joint Photographic Expert Group*
JPG	*See jpeg*
JSON	*Java Script Object Notation*
LP	*Lumbar puncture*
Lx- Ly	*Space between Lumbar vertebrae x and lumber vertebrae y*
mg	*Milligrams*
MoCA	*Montreal Cognitive Assesment*
ml	*Milliliter*
MRI	*Magnetic Resonance Imaging*
MVD	*Multiple Variable Real World Data*
Mx	*Month x*
NCD	*Neurocognitive Disorders*
NPT	*Neuropsychological tests*

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

PBS	*Phosphate Buffered Saline*
PCA	*Principial Component Analysis*
PET	*Positron Emission Tomography*
PM	*Project Manual*
PMH	*Past Medical and Developmental History*
PSG	*Polysomnography*
PST	*Plasma Separator Tube*
PtP	*Paper to Platform*
RBD	*Rapid Eye Movement Behavioural Disorder*
REM	*Rapid Eye Movement*
ROI	*Region of Interest*
ROS	*Review of Systems*
RPM	*Revolutions per minute*
RWD	*Real World Data*
SH	*Social History*
SIMOA	*Single molecule array*
SST	*Serum Separator Tube*
SVD	*Single Variable Real World Data*
TBD	*To Be Decided*
ToC	*Table of Contents*
UI	*User Interface*
UV	*Ultraviolet*
VOSP	Visual Object Space and Perception
WP	*Work Package*
WPL	*Work Package Leader*
XML	*extensible markup language*

Executive Summary

This area is used for the executive summary of the Deliverable. The Multidisciplinary Expert System for the Assessment & Management of Complex Brain Disorders (MES-CoBraD) is an interdisciplinary project combining Real-World Data (RWD) from multiple sources through comprehensive, cost-efficient, and fast protocols towards improving diagnostic accuracy and therapeutic outcomes in people with CoBraD, as represented in Neurocognitive (e.g., Alzheimer's dementia), Sleep, and Seizure (i.e., Epilepsy) disorders. The principles followed within the MES-CoBraD Project can be generalized for the assessment and management of other Chronic Complex Conditions (CCC) by exploiting advanced analytics modules and RWD acquisition protocols.

The MES-CoBraD project identifies seven categories of RWD at the time of this version of the manual:

1. Expert evidence-based questionnaires and structured interview question-trees that inform on people's current clinical symptoms, and medical and social history and phenotypes, derived from information given by patients and their caregivers directly or indirectly
2. Neurological examination and Neuropsychological testing as derived from expert clinician-scientists
3. Neuroimaging, with special focus on Brain Magnetic Resonance Imaging (MRI)
4. Biological samples, with emphasis on deriving cerebrospinal fluid and blood biomarkers
5. Neurophysiological data, especially non-invasive techniques
6. Medical device RWD, and
7. Consumer technology RWD

The current Project Manual (PM) aims to be a continuously updated reference on (a) metrics that are assessed in people with CoBraD belonging to the above categories of RWD, (b) their systematic and standardised quality requirements, and (c) the agreed upon multisource RWD acquisition protocols. The PM reflects the foundation of the project's scientific roadmap by contributing directly to work packages relating to pilot data collection, CoBraD phenotyping and outcome assessment, within the framework of the integrated clinical-research MES-CoBraD Platform. On this basis, it is a reference for users involved in the MES-CoBraD Project.

The PM presents CoBraD, RWD, and their interrelation within the scope of the project by providing a brief overview of CoBraD and their co-morbidity as a justification for the measurement of RWD, while also providing a precise definition of RWD. The PM describes the abstracted methodology for acquisition of each RWD category used by clinical-research partner organisations, and provides information regarding their optimal sources, quality metrics, and storage parameters. In line with the overarching dynamic and modular approach that defines the MES-CoBraD Project, the specific variables and their parameters for each RWD are presented within the respective Appendices of the PM, are regularly updated based on internal partner feedback and project results, as well as review of evolving CoBraD literature. The manual concludes by integrating the various sources and types of RWD into a harmonised and cross-site MES-CoBraD Protocol of RWD acquisition that will serve during the Project's lifetime and, ideally, beyond.

Note: As a user of this manual, anticipate that the PM will be regularly updated, especially with regards to its Appendices, and verify that the version in your hands is the latest one as referred to in the Project's intranet.

The PM in your hands is the first version, reflecting general processes proposed to the European Union that is funding the Project. A second version (i.e., D3.2 - Project Manual – CoBraD RWD, Updated MES-CoBraD Protocols and Quality and Quantity (Q&Q) Evaluation – v2) is planned to be delivered in the three months, integrating feedback from stakeholders across sites, as well as cross-site test run results on the

current protocol that will inform of effective and useful practices to be pursued onwards.

1 INTRODUCTION

Complex brain disorders (CoBraD), as represented in Neurocognitive (e.g., Alzheimer's dementia), Sleep, and Seizure (i.e., Epilepsy) disorders, are chronic conditions that have high prevalence individually and in combination, leading to disability that interferes with activities of daily living and worsens quality of life, increasing mortality risk, and contributing to the socioeconomic burden of patients, their families, and their communities at large. CoBraD share elements of complex pathophysiological processes that lead to a breakdown of brain rhythms and function, further explaining their high comorbidity, and require lifelong medical management that is usually suboptimal without available etiologic thetrapies.[10] The Multidisciplinary Expert System for the Assessment & Management of Complex Brain Disorders (MES-CoBraD) is an interdisciplinary project combining Real-World Data (RWD) from multiple sources through comprehensive, cost-efficient, and fast protocols towards improving diagnostic accuracy and therapeutic outcomes in people with CoBraD and their interdependence. The primary focus of the project is improving the quality of life of patients, their caregivers, and the society at large.

Despite the efforts across the world to effectively improve the assessment and management of CoBraD, several key clinical, research and technological challenges remain that the MES-CoBraD Project aims to address. Most clinical research studies, either observational or clinical trials, in CoBraD are based on concrete sterile conceptual frameworks that are not comprehensive in the assessment of the multidimensional biological and social features of people to determine outcomes, or that fail to assess therapeutic effects in real-world settings. The end result is partial understanding of the pathophysiology of diseases and the failure of several costly clinical trials in CoBraD. Most research is performed in single academic centers, limiting the integration and exploitation of research expertise and resources from across the world. The MES-CoBraD Consortium of expert clinician-scientists, engineers, computer scientists, social health scientists, lawyers, and marketing and communication specialists from across Europe and Israel was created, among many objectives, to provide a novel and impactful unified approach and applicable solutions in improving the comprehensive evaluation and management of CoBraD.

1.1 AIM AND OUTLINE OF THE PROJECT MANUAL

The current Project Manual (PM) aims to be a continuously updated reference on (a) metrics that are assessed in people with CoBraD belonging to the above categories of RWD, (b) their systematic and standardised quality requirements, and (c) the agreed upon multisource RWD acquisition protocols.

The MES-CoBraD project identifies seven categories of RWD: 1) expert evidence-based questionnaires and structured interview question-trees that inform on people's current clinical symptoms, and medical and social history and phenotypes, 2) neurological examination and neuropsychological testing as derived from expert clinician-scientists, 3) neuroimaging, 4) biosamples from sources such as hair, cerebrospinal fluid, and blood, 5) neurophysiology data, 6) medical device RWD, and 7) consumer technology RWD.

The PM presents CoBraD, RWD, and their interrelation within the scope of the project by providing a brief overview of CoBraD and their co-morbidity as a justification for the measurement of RWD, while also providing a precise definition of RWD. The PM describes the abstracted methodology for acquisition of each RWD category used by clinical-research partner organisations, and provides information regarding their optimal sources, quality metrics, and storage parameters. **In line with the overarching dynamic and modular approach that defines the MES-CoBraD Project, the specific variables and their parameters for each RWD are presented within the respective Appendices of the PM, are regularly updated based on internal partner feedback and project results, as well as review of evolving CoBraD literature.** The manual concludes by integrating the various sources and types of RWD into a harmonised and cross-site MES-CoBraD Protocol of RWD acquisition that will serve during the Project's lifetime and, ideally, beyond.

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

1.2 Associated Tasks and Work Packages

The current project manual is a product (deliverable) of the MES-CoBraD Project ***Landscape Analysis and Methodology Work Package*** (WP3). The overall objective of WP3 is defining RWD acquisition and their pre-processing by leveraging the combined expertise and needs of CoBraD stakeholders (to be integrated in the manual at version 2). The tasks and deliverables of WP3 form the foundation for the project's scientific roadmap (figure 1) by contributing directly or indirectly to work packages 4, 5, 6, 7, 8, and 9.

Note: For a complete view of WP3 contribution to the Project and other WP, please review additional deliverables beyond the PM (D3.1 and D3.2).

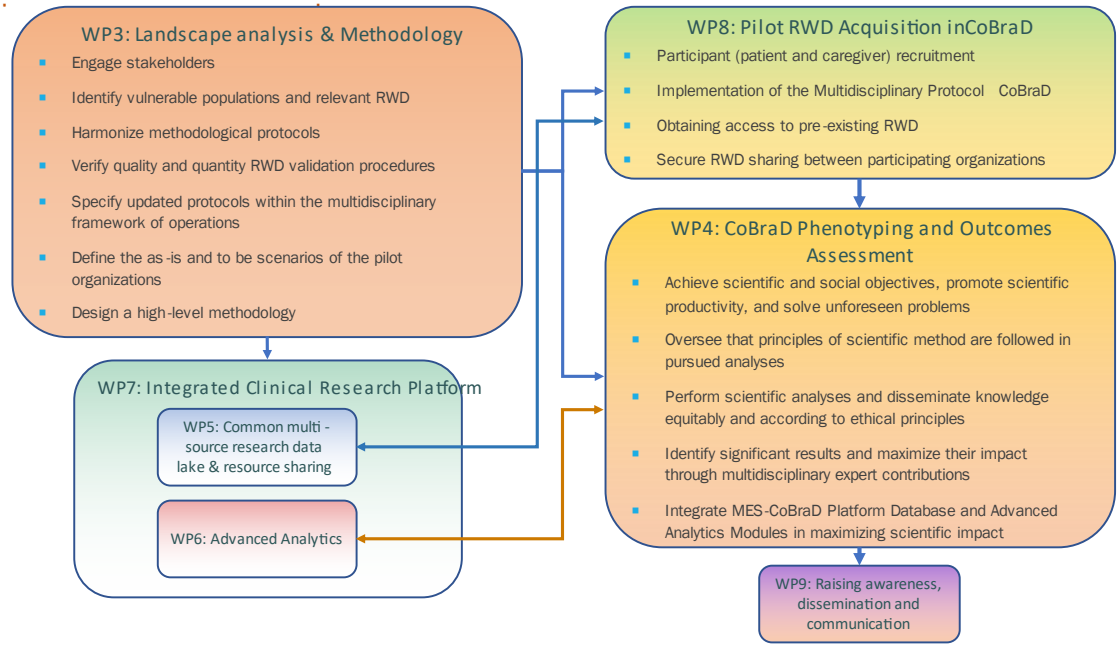


Figure 1: MES-CoBraD simplified scientific roadmap.

1.2.1 Tasks Contributing to the Development of the Current Version of the Project Manual

1.2.1.1 Harmonisation and interoperability protocol of CoBraD RWD acquisition and sharing (Task 3.2)

Through this task, participating organisations define essential information to be extracted from RWD, their parameters, and the integrated protocol for acquiring RWD across sites. Key harmonisation considerations include the specific CoBraD population characteristics and RWD to be acquired during the pilot use cases. To achieve harmonisation, the task identifies necessary hardware, software, and consumables required to be available for or be purchased by pilot partners, as well as localise clinical assessments to regional populations.

Note: Version 2 of the PM with further integrate results from Task 3.1 on information acquired through Landscape Analysis and Stakeholder Engagement activities to incorporate feedback from regional stakeholders, as well as test run results.

1.2.2 Work Packages that Directly Utilise the Project's Manual

1.2.2.1 Pilot RWD acquisition in CoBraD (Work Package 8)

Through WP8, pilot RWD will be acquired in accordance to the protocol of the PM. RWD of participants with CoBraD will be securely stored post-anonymisation in a central database developed through WP5 for secure sharing between the consortium partners and, eventually, the public.

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

1.2.2.2 CoBraD phenotyping and outcome assessment (Work Package 4)

This WP defines the methodological process for analysing RWD acquired during pilot data acquisition (WP8) to address the challenges faced in CoBraD, and sharing the results within the consortium and through scientific publications prior to wider dissemination. Any planned analyses in WP4 depends on the PM, which defines how RWD are structured, harmonised, and collected through a predefined protocol.

1.2.2.3 Integrated Clinical Research platform (Work Package 7)

WP7 will allow feature visualisation, maps, class activation mapping, and sensitivity analyses, where certain parts of the image are hidden to the effect on prediction. Going one step forward, graphical tools will be employed to visualise in an efficient manner all the significant interrelationships between hand-crafted features and deep learning ones that are developed through WP6, while leveraging the backend infrastructure developed in WP5. The main aim of WP7 within the scope of WP3 is to prototype the MES-CoBraD Platform through technological components that support Data Collection according to the MES-CoBraD Protocol.

1.2.2.4 Common multi source research data lake & resource sharing (Work Package 5)

The aim of the WP is to provide the platform (WP7) with a data source layer taking care of three specific aspects: the acquisition from heterogeneous data sources, the data anonymisation and the sharing of data and their relative analysis algorithms. In order for the WP5 to acquire data from multiple source it requires the data structure conceptualisation provided by WP3, specifically as noted in the current manual.

1.2.2.5 Advanced analytics (Work Package 6)

This WP provides tools needed to accomplish scientific research in the form of modules. Artificial Intelligence (AI) and machine learning tasks will develop the AI knowledge-based system, enabling machine learning capabilities and emulation of human cognitive functions. The structure, context, and interrelatedness of RWD as explicit in the manual WP3, will be central to their analysis.

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

2 TYPES OF COMPLEX BRAIN DISORDERS, THEIR CLINICAL EVALUATION, DISABILITY AND SOCIETAL IMPACT

2.1 MAJOR NEUROCOGNITIVE DISORDERS

Major neurocognitive disorders (NCD), historically called dementia, are a set of syndromes in which people have cognitive impairment (i.e., deficits in memory, language, visuospatial skills or executive function) that interferes with their ability in pursuing daily activities compared to the past. If symptoms are milder and cognitive impairment does not interfere with daily activities, they are called minor neurocognitive disorders (minor NCD) or mild cognitive impairment. The different NCD syndromes are defined according to their specific cognitive and behavioural symptoms and their severity. Each dementia syndrome is associated with one or more underlying causes with neurodegenerative brain diseases, such as Alzheimer's disease (AD), and vascular brain diseases being the most common.

Characteristic features of neurodegenerative brain diseases are the abnormal clustering of brain cell proteins, called a proteinopathy, and brain cell death. Both features tend to precede NCD symptoms by more than a decade. A specific neuropathological diagnosis is dependent on the proteinopathy that is observed under the microscope, such as amyloid and tau in AD, α-synuclein in Lewy Body Disease, and tau and TDP-43 in Frontotemporal Lobar Degeneration. Proteinopathy is only one aspect of neurodegeneration, and several processes, including immune, vascular, metabolic, and genetic, contribute to gradual functional and structural changes.

Vascular brain disease, on the other hand, refers to brain damage due to poor blood perfusion as a result of occluded vessels. This occlusion can be caused by a clot, as in atrial fibrillation or carotid artery disease, or direct damage to the brain vessels from vascular risk factors (e.g., hypertension, hyperlipidemia, diabetes, smoking, sleep apnea). When there is damage to critical brain areas, NCD develop.

Diagnosis of NCD by professional guidelines requires, in addition to a clinical evaluation by a trained specialist, neuropsychological evaluation through paper and pencil or computer-based cognitive tests, brain imaging (preferably magnetic resonance imaging [MRI]), and blood or cerebrospinal fluid testing, explaining the inclusion of the respective types of RWD in the MES-CoBraD Protocol. NCD have high morbidity and mortality, with people eventually requiring daily assistance that increases the direct and indirect cost of care. The high disability and prevalence of 5-8% for people over 60 in NCD, especially for low socioeconomic status communities, places an immense societal burden as the world population ages and resources for optimal evaluation are scarce. The exact global financial cost for dementia care is unknown, with direct costs exceeding 232 billion dollars in the United States, and indirect costs to unpaid family caregivers providing 83% of care being hard to establish.

2.2 SLEEP DISORDERS

Primary sleep and circadian sleep rhythm disorders, collectively referred here as sleep disorders, are a set of syndromes that interfere with people's sleep quality or quantity (primary sleep disorders), or timing (circadian sleep rhythm disorders), and lead to worse quality of life.

Sleep disorders interfering with sleep quality are sleep-related breathing disorders, such as sleep apnea, where people do not breathe effectively during sleep, parasomnias, such as REM sleep behaviour disorder, in which abnormal behaviours surface during different sleep stages, and sleep-related movement disorders, in which people have excessive movements around or during sleep periods. Sleep disorders of poor sleep quantity are insomnias, in which people are unable to fall asleep, maintain sleep continuity, or wake up earlier than desired despite given the opportunity to sleep, and central hypersomnia, such as narcolepsy, in which people have excessive somnolence that is not caused by poor

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

sleep quality and who tend to sleep if given the opportunity.

Sleep disorders interfering with sleep timing are circadian sleep rhythm disorders, in which people may have good sleep quality and quantity, but the timing of their sleep periods interferes with social demands, such as sleeping too early in the evening (advanced sleep phase disorder) or too late (delayed sleep phase disorder), or having an inconsistent or irregular sleep-wake schedule across the 24-hour cycle.

The pathophysiology of sleep disorders varies depending on the syndrome at hand. In most cases, central nervous system dysfunction is observed in syndromes of hypersomnia, irregular sleep-wake rhythm disorder, certain insomnias associated with neurological diseases,, parasomnias, and even sleep-disordered breathing. Mental and physiological hyperarousal is a key mechanism of insomnia, justifying cognitive behavioural therapy for insomnia as a first line treatment before medications. Finally, aging, narrow upper airway anatomy, and male sex are the main risk factors associated with obstructive sleep apnea, in which the upper airway becomes narrower or collapses during sleep, leading to hypopneas or apneas that lead to arousals or hypoxias, thus interfering with sleep quality.

A comprehensive sleep evaluation may require, in addition to a clinical evaluation by a specialist, a sleep study at night (polysomnography [PSG]) or in the day (multiple sleep latency test or maintenance of wakefulness test), during which brain activity, respiratory function, muscle movement, heart rate, and even video monitoring are pursued in diagnosing sleep-related breathing disorders, parasomnias and hypersomnia. Activity measurements over several days with a special wrist accelerometer, called actigraphy, can help establish insomnia or circadian sleep rhythm disorders. Specific bio sample tests help in diagnosing certain syndromes, such as narcolepsy. The above explain the choice of the respective RWD as integral to the MES-CoBraD Protocol.

Sleep disorders are extremely common where, beyond typically causing drowsiness, impaired attention, and poor mood, they further increase occupational hazard, including worse work performance and increased risk for motor vehicle accidents. Moreover sleep disorders pose an increased risk for medical comorbidities such as: diabetes, obesity, cardiovascular disorders, psychiatric conditions, neurocognitive disorders, and earlier all-cause mortality. Despite a combined prevalence of one in four to one in seven people having a sleep disorder at any given time, few receive care or even know of having a sleep disorder. Similar to other CoBraD, sleep disorders are more common in low socioeconomic strata, where there is poor access to specialists and to effective therapies. The above explain the high cost of care for sleep disorders reaching close to 1.55% of the gross domestic product of high-income countries.

2.3 SEIZURES AND EPILEPSY

Epilepsy or seizure disorders are a group of syndromes that are defined by an enduring predisposition for recurrent unprovoked epileptic seizures, defined as abnormal hypersynchronous and sustained neuronal excitability. Epilepsy leads to adverse neurobiological, cognitive, psychological, and social effects. Aberrant neuronal firing leads to temporary abnormal brain function which can present as hyper motor convulsion, subjective abnormal perceptions (e.g., gustatory, olfactory), cognitive or behavioural impairment, or decreased level of alertness and interaction with the environment. The pathophysiology of seizure disorders is varied, including malformations during brain development, neurodegeneration, trauma, brain ischemia, tumors, infections and metabolic derangements. In half of people the cause of epilepsy is unknown, explaining seizure syndromes categorised as structural, genetic, infectious, metabolic, immune and unknown.

A comprehensive seizure disorder evaluation requires, a clinical evaluation by a specialist and electroencephalograpy (EEG), during which electrical brain activity is recorded with or without video monitoring to establish the presumed abnormal neuronal excitability. Blood or cerebrospinal fluid tests help in diagnosing certain genetic, autoimmune or metabolic syndromes. Brain imaging and neuropsychological testing are required for almost all patients, to establish the neurobiological and

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

cognitive-behavioural consequences of seizure disorders. The above justify the inclusion of the respective RWD types in the MES-CoBraD harmonised Protocol.

The social burden of seizure disorders is a combination of its high prevalence and morbidity, with an estimated 50 million people worldwide having epilepsy at any given point and an estimated five million people being diagnosed with epilepsy annually. The social and economic implications of seizure disorders are further complicated by the associated social stigma. The cognitive-behavioural deficits associated with seizure disorders further interfere with a person's work productivity and daily quality of life, whereas the complex morbidity with frequent hospitalisations and visits to the emergency department places a disproportionate burden to healthcare systems. The annual direct cost for epilepsy in the United States is approximately $8,412 to $11,354 per patient and indirect costs range between 12 – 85% of direct costs. The Global Burden of Disease study 2015 ranks epilepsy as the 5th most burdensome neurologic disorder worldwide in terms of disability-adjusted life years.

2.4 CoBraD Comorbidity

In the framework of CoBraD, research results by health participating organisations (POs) in this proposal and other groups prove the high comorbidity and overlapping complex pathophysiology of NCD, sleep, and seizure disorders. As shown in Table 1, any of the three CoBraD may precede or follow another CoBraD, and by an average of four years in many cases, suggesting common pathophysiological pathways between CoBraD, whereas their comorbidity accelerates CoBraD disability and worsens prognosis.

Table 1: Prevalence of individual CoBraD in the general population (diagonal cells) and the prevalence of a secondary CoBraD when a primary CoBraD is previously diagnosed.

		Primary CoBraD		
		Neurocognitive Disorders	Sleep and circadian rhythm disorders	Epilepsy and seizure disorders
Secondary CoBraD	Neurocognitive Disorders	2-7%	27-38%	16-40%
	Sleep and circadian rhythm disorders	70-80%	14-25%	25-75%
	Epilepsy and seizure disorders	11-40%	Unknown	1-3%

A large body of research studies pursuing CoBraD deep-phenotyping and endo-phenotyping (i.e., the precise and comprehensive analysis of a syndrome's symptoms and its underlying brain processes), verify common pathological processes between CoBraD syndromes. CoBraD deep-phenotyping studies using precision medicine protocols in clinical populations revealed the actual comorbidity of CoBraD can be three times higher than suggested in epidemiological studies. Thus, the presence of a primary CoBraD is highly predictive of eventual emergence of a second CoBraD over time, as in the case of Rapid Eye Movement Behavioural Disorder (RBD), whose presence precedes cognitive-motor symptoms of synucleinopathies (e.g., Lewy Body Disease) by an average of eight years.

2.5 Sex and Gender Considerations

The MES-CoBraD Project takes special attention to understanding sex and gender factors in CoBraD, as well as in the conduction of research within the Consortium. In the project, the role of sex and gender is taken into account through "intersectionality," so that appraisals can be advanced on multi

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

characteristics of social identity beyond sex and gender, even if only in a counterfactual way. The theory of intersectionality argues that various forms of discrimination centered on race, gender, class, disability, sexuality, and other forms of identity, do not work independently but interact to produce particularized forms of social oppression.

In order to perform this analysis, we can envisage the following RWD and their sources.

- RWD obtained through pre-existing databases,or acquired through questionnaires and question-trees addressed to people with CoBraD (see Chapter 4.1 below). In addition to RWD from common categories of (a) men and women, (b) education, and (c) socioeconomic status will be acquired, further detailed questions will be available on delving into sex and gender analysis
- Consortium-targeted longitudinal surveys on sex and gender metrics within and between teams.

2.6 CLINICAL SYMPRTOMS AND HISTORY

Clinical Symptoms and History (CSH) RWD are the cornerstone RWD of any assessment, and represent individuals' salient complaints related to CoBraD, e.g memory problems in the setting of NCD, as well as historical clinical and social information that contribute to or are affected by symptoms. Symptom relief is also one of the most important benchmarks for a successful therapy. Due to the nature of CoBraD, multiple symptoms may occur in a single person, may manifest with variable severity, and differentially impact quality of life. Moreover, symptoms may be caused by other comorbid medical conditions or their therapies. The high level of complexity and dimensionality of CSH requires systematic deep phenotyping of symptoms and history, ideally by trained experts as per current state-of-the-art, and subsequently integrated with findings on physical exam and laboratory testing. The CSH can be provided directly by patients or their caregivers, or indirectly by clinical-research staff, and represent:

1. Troubling symptoms of a patient with regards to their History of Present Illness (HPI) following a Chief Complaint (CC), current Review of Systems (ROS), and current Functional Status (FS), as well as
2. Historical information of Past Medical and Developmental History (PMH), Social History (SH), Family History (FH), and Medications

This type of information allows clinicians and researchers to identify features in a person's symptoms and history that allow deep-phenotyping of a patient's medical and social condition, including a possible underlying pathophysiological process, and guide further workup towards accurate diagnoses and treatment choices.

2.6.1 SOURCE, ACQUISITION METHOD, DATA ABSTRACTION

There are four main sources from where CSH are derived:

1. Patients
2. Caregivers
3. Clinicians-researchers and clinical-research coordinators
4. Existing databases

Most of the CSH RWD are acquired directly from patients or their caregivers, whereas certain information reflect metrics entered by clinician-scientists or clinical-research coordinators after interviewing patient-caregiver dyads. CSH RWD only reflect patient-caregiver perceptions and not clinician-scientist or coordinator assessments during an encounter with patients-caregivers, which are represented through

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Physical Examination & Neuropsychological Testing RWD. Under this premise, clinical-scientist staff serve as facilitators and, accordingly, mediators in CSH RWD acquisition.

There are **three acquisition formats** for CSH RWD:

5. Structured questionnaires
6. Decision tree specialty-specific questions
7. Database extraction (e.g., Electronic Medical Records [EMR])

All formats can represent the same underlying information, so that a question in a decision tree can also be a question of a questionnaire, or be represented as a variable in a database. The high-level RWD semantic associations that MES-CoBraD is built to account for allows Advanced Analytics modules easier data integration and analysis, that are subsequently exploited through its Expert System.

Structured questionnaires are series of predefined questions provided in sequence to a patient or their caregiver to answer according to specific instructions. In many cases, a questionnaire score is computed through a weighted sum of all questionnaire answers, representing CSH severity at a given time, and, thus, facilitating follow up of CSH over time. This also explains why questionnaires are almost always used in clinical research and trials as primary or secondary endpoints. The MES-CoBraD Platform allows a clinician-scientist to choose through their protocols one or more structured questionnaires, or have derived questionnaire scores from answers provided through specialty-specific question trees or database information as described below. Most questionnaires represent in one form or another the information obtained through ROS, FS, PMH, FH, SH, and Medications.

When it comes to the reason a person visited the clinic in the first place, however, exhaustively going through arbitrary questionnaires is not practical, and can be misleading. Instead, when a patient visits their clinician, they provide CSH RWD according to **question trees**, where the answer of one question leads to deciding which is the best next question. This is best represented in a clinician's HPI questions following a person's answer to a Chief Complaint (CC) for visiting a clinic. These series of questions can be identical to the individual elements of one or more questionnaires, but they are asked (a) only if pertinent according to previous questions, and (b) often relate more to the CC. Nonetheless, if a patient or their caregiver responds to question trees that include all questions of a questionnaire, an estimated score of that questionnaire can also be derived.

Note that different CSH RWD acquisition formats can lead to different responses by users, however, as long as the question is understood by the user (patient, caregiver, or clinician-scientist), these differences are small in most cases.

Data abstraction of CSH RWD is based on three orthogonal categories, each having a hierarchical structure:

8. CSH-based RWD
9. STS (Severity, Temporal, Spatial)
10. CCC-specific RWD (i.e., CoBraD-specific RWD)

The orthogonal nature of the above categories allows for RWD to be classified on several dimensions, thus facilitating the grouping of variables and identifying semantic associations between them for future analyses by clinician-scientists and Expert Systems.

All three data abstraction categories and their hierarchical structure are expanded on in Appendix 1

The **CSH-based RWD** category structure is fixed for all types of CCC, including CoBraD, and reflects principles of the clinician's interview, organised as subcategories.

Such a categorisation allows for deciding on the sequence and grouping of presenting CSH RWD to users,

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

verifying that CSH-based RWD are presented and completed by a user, thus optimising completeness of data acquisition in clinic or research.

The **STS** category, in a self-explanatory manner, indicates if CSH RWD information refers to presence/absence and severity of CSH RWD (i.e., Severity), whether it refers to the duration, periodicity, and speed of symptom progression (i.e., Temporal), or the spatial features of these symptoms (i.e., Spatial). Note that presence of RWD that reflect Temporal and Spatial features are conditional on presence of RWD Severity.

The **CCC-specific** RWD category data abstraction is variable according to the specific CCC a project aims to address. For the MES-CoBraD pilot, hierarchical data abstraction of CoBraD represents CSH relevant to neurocognitive, epilepsy, and sleep disorders. In line with the conceptual framework of the MES-CoBraD Project, the CCC-specific category domains (i.e., hierarchal data abstraction) are modular in nature, allowing tailoring to varied clinical-research settings and new evolving scientific knowledge. The level of depth within this hierarchical structure between superordinate to subordinate levels is similarly modifiable and varies according to the needs of the clinicians and scientists. Lower levels of abstraction (subordinate) allow for finer-grained information description of CSH RWD.

Note that CCC-specific category organisation corresponds to CCC-specific hierarchical data abstraction of clinical examinations and laboratory tests applied in clinical-research practice to objectively examine these CCC-specific categories. For example, neuropsychological tests reflect metrics of cognition, whereas sleep studies metrics of sleep. To that end, there is a large overlap of the respective data abstraction between categories of RWD (see respective chapters of the MES-CoBraD Manual).

For example, a CSH variable relating to “seizure frequency” can be grouped in three dimensions as HPI (CSH-based RWD), Severity, and Frequency and Rhythmicity (STS), and Epilepsy > Seizure (CCC-specific RWD).

2.6.2 QUALITY CONTROL

Quality control of CSH RWD reflects

11. Responses within a prespecified value range
12. Completeness of data within protocol requirements
13. Correspondence of user responses to the intended construct

Responses within a prespecified value range are easily controlled for DtP acquisitions where the acceptable range of responses is provided in the Appendix table with the respective variable’s elements.

Instead, errors are more likely to occur in PtP acquisitions, where participants may (a) circle multiple responses on a sheet of paper (e.g., “4” and “5” when they feel a value of “4.5” is more representative; or “4,” “5,” and “6” when they experience more variability in their symptom perception), or respond with value types different from the question at hand (e.g., “a lot” as an answer to “What year did your symptoms start?”) often reflecting poor comprehension. These types of errors will be mitigated by having clinical or research coordinators review together with participants their answers, targeting specific CSH RWD where paper responses are invalid and providing clarifications. If an invalid response is missed, then the acceptable response will be the median or mean value (depending on data type). If this is not possible, then the response is considered invalid, and the respective code value (e.g., -99) can be entered instead. Although most CSH RWD variables have the same missing or erroneous value codes (e.g., -99 or -88), these are provided for each variable within the respective Appendix.

Completeness of data within protocol requirements

Clinical or research protocols require that participants provide a set of information to allow diagnosis or direct treatment decisions. Completeness of data falls under a protocol’s RWD quantity requirements and

The MES-CoBraD project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 965422

as RWD is entered on the Platform, coordination tools can inform of missing information to be acquired.

Additionally, unique to CSH RWD (as well as certain clinical and neuropsychological examination RWD), information is acquired through a sequence of questions when the acquisition format is a questionnaire or a question tree (see above). In the case of DtP acquisition, the Platform itself is able to direct the users to answer unanswered mandatory questions before the protocol proceeds with the next question. In the case of PtP acquisition, a clinical-research coordinator or a clinician will have to ask the user (usually a patient or a caregiver) to complete the unanswered questions.

Correspondence of user responses to the intended construct

In addition to quality control of RWD input, an often-overlooked component of quality control is whether CSH RWD represent the intended construct, i.e., information. To mitigate this, all questions through which CSH RWD are acquired are vetted through pilot organisations, and translated and localised to their regional populations and assessed during pilot protocol test runs. Furthermore, participants will be regularly assessed through clinicians on their responses, and over time (either during protocol test runs or pilot period) poor construct representations will be flagged and corrected. Finally, construct validity of individual or composite variables can also be pursued through Platform-based analyses.

2.6.3 DATA STORAGE

CSH RWD entered directly in an anonymised manner DtP will be available to the Platform, and at the same time will be available to the local site's secure and GDPR compliant repository.CSH RWD entered via PtD, will have paper-based responses scanned in pdf format on the local site's secure and GDPR compliant repository, and clinical-research coordinators will enter the information in an asynchronous manner to the Platform.

2.7 ETHICAL CONSIDERATIONS

The Consortium is fully committed to adhere to the highest ethical, fundamental rights and legal standards, as recognised at the European Union and International levels, including the *Charter of Fundamental Rights of the EU* (2000/c 364/01), the *Clinical trials Regulation* (EU 536/2014), the *General Data Protection Regulation* (GDPR) (Regulation (EU 2016/679) and *The European Code of Conduct for Research Integrity* (ALLEA, 2017 revised edition), and the *OECD Council Recommendations on Health Data Governance*. Moreover, the project will be carried out in accordance with the Declaration of Helsinki and Taipei as well as the Convention on the Protection of Human Rights and Human Dignity in Biology and Medicine.

Any and all research activities will be conducted based on the following core medical ethics principles:

- **Autonomy**: one should respect the right of individuals to make their own decisions
- **Non-maleficence**: one should avoid causing harm
- **Beneficence**: one should take positive steps to help others
- **Justice**: benefits and risks should be fairly balanced.

Clinicians, researchers, and any member of the Consortium involved in the project are subject to these standards and are required to be compliant with the following principles:

- **Reliability** in ensuring the quality of clinical practice and research, reflected in the design, the methodology, the analysis and the use of resources
- **Honesty** in developing, undertaking, reviewing, reporting and communicating clinical and research information in a transparent, fair, full, and unbiased way

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

- **Respect** for colleagues, clinical and research participants, society, ecosystems, cultural heritage, and the environment
- **Accountability** for the research from idea to publication, for its management and organisation, for training, supervision and mentoring, and for its wider impact

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

3 REAL WORLD DATA STRUCTURE

The focus of this chapter will be on the Real World Data (RWD) coming from the clinical facilities. The aim here is to provide terminology and technical considerations that are relevant for the MES-CoBraD project.

The RWD that adhere to the guidelines described hereafter can be considered "***MES-CoBraD Ready***" since they will be easily managed and elaborated in automatic way.

There are **two User-Interface (UI) methods** for acquiring Clinical Symptoms and History (CSH) RWD:

1. Direct-to-Platform (DtP)
2. Paper-to-Platform (PtP)

It is well established that there are fewer mistakes and omissions through DtP RWD acquisition, such as limiting answers to acceptable values or mandating answers be completed to proceed. Nonetheless, in many settings this is not feasible, especially when considering vulnerable populations with little electronic education and experience, when resources to electronic acquisition tools (computers, tablets, smart-phones etc.) are limited, or, even, when electronic acquisition tools fail (e.g., programs crashing, drained batteries). In such cases, PtP acquisitions are required as alternate or backup methods.

Although either method can bypass the mediation of clinical-research staff during CSH RWD acquisition, in many cases staff are required to clarify questions, especially as novel protocols are being developed and streamlined. This may be achieved asynchronously, however, after patients and caregivers have provided the respective CSH RWD, and following Protocol guidelines to optimise staff time and cost. In the rest of the chapter the focus will be more on the DtP acquisition method.

As the chapter 4 of this document will show, during the execution of the MES-CoBraD project several types of RWD will be collected and managed.

Each of these RWD types has its own optimal representation format. For example, a patient's blood pressure over time can be well represented through a series of data points indexed in time order; this kind of representation, named a time series, is optimal also for several other cases like (but not only for) other biological levels.

But not always, the focus will be on the evolution over time of a single value; for example, in the case of an MRI, the interest is not on the observation of a single value but on the observation of an image that represents the current status of the patient under several point of view at the same time.

In still other cases, like for example anamnesis questionnaires, the optimal representation will be a more complex object-like one.

From a technological perspective, this means that the MES-CoBraD project must will be able to deal with a plethora of different input types. Moreover each RWD type can be serialised in several valid data formats: e.g. a time series can be serialised through a simple CSV file or through a JSON Array.

The following table shows two possible (and perfectly valid) representations of a generic time series.

Table 2 Two valid representations of a sample time-series.

CSV	***JSON***
"DateTime", "Value" *"01/04/2021", "14.5"* *"02/04/2021", "18.5"* *"03/04/2021", "20.3"* *…*	*[{* *"DateTime": "01/04/2021",* *"Value": "14.5"* *},{* *"DateTime": "02/04/2021",*

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

	"Value": "18.5" *},{* *"DateTime": "03/04/2021",* *"Value": "20.3"* *},...]*

The goal of this section is to define a general framework for the RWD serialisation that can act as a guideline for the RWD collector.

The expected RWD types to collect is the one depicted in the following image

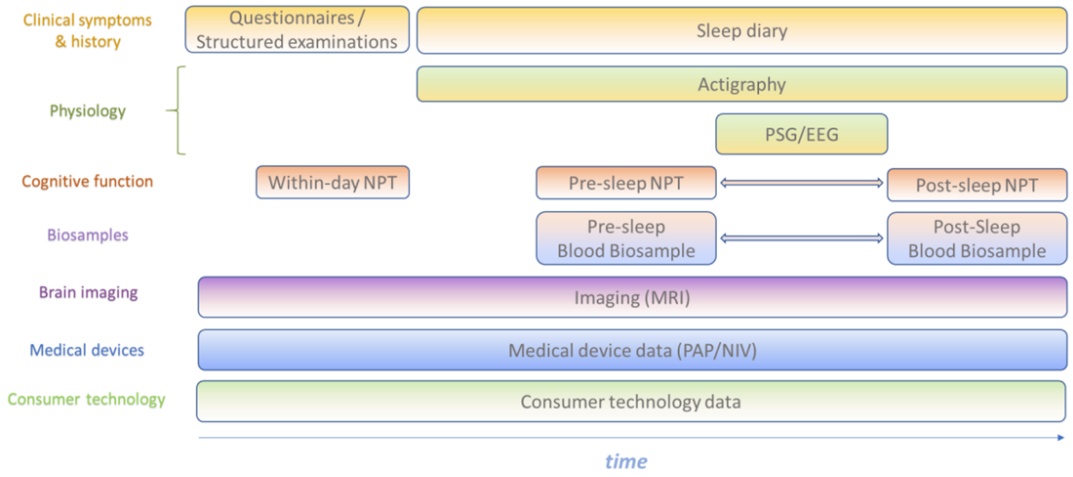


Figure 2 MES-CoBraD Multidisciplinary Multisource RWD Assessment Protocol. The above figure expected types of RWD types.

To facilitate RWD acquisition towards usable information, we reduced RWD into categories. Accordingly, in addition to their semantic information, RWD can be categorised in the following three macro-areas:

1. Single Variable RWD (SVD - e.g., single value data, etc.)
2. Multiple Variable RWD (MVD - e.g., questionnaires, etc.)
3. Complex Structured RWD (CSD - e.g., images, free text, etc.)

The following sections address each RWD type, presenting how the proposed framework intends to address category specific features.

Each variable is characterised by a set of parameters; such set is deeply described in the Appendix.

3.1 SINGLE VARIABLE RWD – SVD

Single Variable RWD (SVD) are data that contain just one semantic level.

The evolution over time of a SVD can be managed like a time series. The MES-CoBraD Project is able to collect those data according to the following prioritised channels and data formats:

1. Direct connection to remote portals' Application Programming Interface (APIs)
2. Direct connection to structured databases
3. Collection of the data in CSV format
4. Direct connection to sensors and devices

The preferred way of collecting the SVD is the direct **connection to remote portals' APIs**. In this case, the only constraint is that the response of the APIs must be structured in some way and available in a non-proprietary open format. Valid formats are JSON , XML and CSV .

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

The second way of collecting SVDs is the direct **connection to databases** containing the raw data in a structured way. In this case, no specific constraint on data is necessary since the database itself will ensure that the data have a machine-readable structure.

In case, for whichever reason, the first two strategies are not suitable, the third supported approach is the periodic collection of the data in **CSV format**. The validation of the file will be done at two different syntax levels:

1. the first validation level will be done against the guideline of the official RFC4180 specification to check that the CSV is generally well-structured;
2. once the validation of the CSV is ensured, the following additional validation will be applied:
    a. column separator check: the column separator must be the comma sign (,) and it must be consistent across the entire file

Table 3 Valid and invalid use of column separator check

Valid Entry	Invalid Entries
“1997-07-16T19:20:30+01:00”,“24”	“1997-07-16T19:20:30+01:00”;“24” “1997-07-16T19:20:30+01:00”:“24” “1997-07-16T19:20:30+01:00”|“24”

- **quotation check**: the CSV values must be quoted through double quotes (“)

Table 4 Valid and valid use of quotation check

Valid Entry	Invalid Entries
“1997-07-16T19:20:30+01:00”,“24”	‘1997-07-16T19:20:30+01:00’,’24’ 1997-07-16T19:20:30+01:00,24

- **schema check**: the overall CSV file must have one field named **datetime** – containing the value sample date and time – and one field named **value** – containing the variable value. No line containing less or more columns must be present within the file

Table 5 Valid and invalid csv name for single variable data.

Valid Entries	Invalid Entries
“datetime”,“value” “1997-07-16T19:20:30+01:00”,“24” “1997-07-16T19:20:30+01:00”,	“datetime”,”value” “1997-07-16T19:20:30+01:00” “1997-07-16T19:20:30+01:00”,“24”,”…”
	“time”,”val” “1997-07-16T19:20:30+01:00”,“24”

- **datetime column check**: the dates must be compliant with the YYYY-MM-DDThh:mm:ssTZD pattern of the ISO8601 format .

Table 6 Valid and invalid methods of noting date and time in single variable data

Valid Entry	Invalid Entries
“1997-07-16T19:20:30+01:00” “1997-07-16T00:00:00+00:00”	“1997-07-16T19:20:30” “1997-07-16T19:20” “1997-07-16” “1997/07/16T19:20:30+01:00” “16/07/1997T19:20:30+01:00”

- **value column check**: the value type should be consistent across the entire file. The first available value determines the type of the expected data in the rest of the file. Values like “NULL”, “Unknown”, “N/A” or similar will be treated as strings, so in case of unavailable data

The MES-CoBraD project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 965422

the value field can be left blank.

Table 7 Valid and Invalid methods of value denotation in single variable data

Valid Data	Invalid Data
“datetime”,”value” “1997-07-16T19:20:30+01:00”,“24” “1997-07-16T00:00:00+00:00”,“24”	“datetime”,”value” “1997-07-16T19:20:30+01:00”,“24” “1997-07-16T00:00:00+00:00”,“NO”
“datetime”,”value” “1997-07-16T19:20:30+01:00”,“YES” “1997-07-16T00:00:00+00:00”,“NO”	“datetime”,”value” “1997-07-16T19:20:30+01:00”,“24” “1997-07-16T00:00:00+00:00”,“Unavailable”
“datetime”,”value” “1997-07-16T19:20:30+01:00”,“24” “1997-07-16T00:00:00+00:00”,	“datetime”,”value” “1997-07-16T19:20:30+01:00”,“24” “1997-07-16T00:00:00+00:00”,“”

The last option to enter data in the MES-CoBraD Platform is through **direct connection to sensors and devices** that collect data. In this case, no specific constraint on data is necessary, since the sensors and devices themselves will ensure that the data have a machine-readable structure.

3.2 MULTIPLE VARIABLES RWD – MVD

Multiple Variable RWD (MVD) are those data that contain more than one semantic level. According to this definition, the questionnaires submitted to patients can be considered a particular type of MVD.

The evolution over time of a MVD can be managed as a multidimensional time series. The MES-CoBraD Project is able to collect such data according to the following prioritised channels and data formats:

1. Excel file (only valid for Questionnaires)
2. Direct connection to remote portals’ APIs
3. Direct connection to structured databases
4. Collection of data in CSV format
5. Direct connection to sensors and devices

In the particular case of Questionnaires, the accepted data format is a simple Excel file. Each Excel file should adhere to the following guidelines:

- Cell styling is generally accepted, but the styling must not be used to carry any semantic information.

Table 8 Valid and invalid use of excel cell styles.

Valid Example	Invalid Example
Responses like “YES” or “NO” where the cells are styled in green or red according to the text value (just for visualisation purposes)	A red empty cell means “NO”, a green empty cell means “YES”.

- No Excel proprietary functionalities must be enabled in the file. Formulas, filters and tables are not allowed.
- In case of multiple sheets in a single file, all sheets should contain data for entry into the Platform.
- Each Questionnaire type must have a reference documentation uploaded in the MES-CoBraD Project Shared Document Library that is intended to present the structure of the excel file itself.

The MES-CoBraD project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 965422

All the MVDs, other than questionnaires, adhere to the same collecting strategies as SVDs. The preferred strategy is the direct **connection to remote portals' APIs**; the only constraint being that the response of the APIs is structured in some way and available in a non-proprietary open format. Valid formats are JSON, XML and CSV.

The second way of collecting MVDs is the direct **connection to databases** containing the raw data in a structured way. In this first case, no specific constraint on data is necessary since the database itself will ensure that the data have a machine-readable structure.

In case, for whichever reason, the first two strategies are not suitable, another supported approach is the periodic collection of the data in **CSV format**. The validation of the file will be done at two different syntax levels:

The first validation level will be done against the guideline of the official RFC4180 to check that the CSV is generally well-structured; once the validation of the CSV is ensured, the following further validation will be applied:

- column separator check: the column separator must be the comma sign (,) and it must be consistent across the entire file , as in table 3-2)
- quotation check: the CSV values must be quoted through double quotes (“), as in table 3-3.
- schema check: the overall CSV file must have:
  - one datetime column, containing the value sample date and time
  - one column for each supported variable, containing the corresponding value; the column name must be equal to the variable name itself.

No line containing a different amount of columns must be present within the same file.

Table 9 Valid and invalid csv name for multiple variable data.

Valid Entries	Invalid Entries
“datetime”,“rem_duration”,“sleep_start” “1997-07-16T19:20:30+01:00”,“4”,“1997-07-15T22:18:10+01:00”	“datetime”,“rem_duration”,“sleep_start” “1997-07-16T19:20:30+01:00” “1997-07-16T19:20:30+01:00”,“4”

- datetime column check: the dates must be compliant with the YYYY-MM-DDThh:mm:ssTZD pattern of the ISO8601 format. (see Table 9)
- value columns check: the values of each column should be consistent across the entire file. The first available value determines the type of the expected data in the rest of column. Values like “NULL”, “Unknown”, “N/A” or similar will be treated as strings, so in case of unavailable data the value field can be left blank

Table 10 Valid and invalid denotation of multiple variable data.

Valid Data	Invalid Data
“datetime”,”rem_duration” “1997-07-16T19:20:30+01:00”,“8” “1997-07-17T00:00:00+00:00”,“7.5” “1997-07-18T00:00:00+00:00”,	“datetime”,”rem_duration” “1997-07-16T19:20:30+01:00”,“8” “1997-07-17T00:00:00+00:00”,“7.5h” “1997-07-18T00:00:00+00:00”, “180m”
“datetime”,”good_sleep_quality” “1997-07-16T00:00:00+00:00”,“YES” “1997-07-16T00:00:00+00:00”,“NO”	“datetime”,”rem_duration” “1997-07-16T19:20:30+01:00”,“8” “1997-07-16T00:00:00+00:00”,“ N/A”

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

3.3 COMPLEX STRUCTURED RWD – CSD

Complex Structured RWD (CSD) are those data that, differently from SVDs and MVDs, are not easy to serialise in plain formats like CSV or JSON. Some examples of such data are EEG and MRI, but also complex, nested and object-like data structures. The following sub-sections describe the framework guidelines for collecting these types of data.

3.3.1 IMAGES

In order to be collected in the MES-CoBraD platform, the images must adhere to the following rules:

- be in JPG format;
- the filename adhere to the following pattern:

  <subjectID>_<dataType>.<extension>

  where

  - <subjectID> unique anonymised identifier of participant entirely written in lowercase
  - <dataType> is the data type the image represents entirely written in lowercase
  - <extension> is the file extension; accepted extensions are .jpg .jpeg .jpe .jif .jfif .jfi (case insensitive).

Table 11 Valid and invalid filenames for images

Valid filename	Invalid filename
eng2021062301_mri.jpg	antoninosirchia_MRI.jpg
eng2021062301_mri.JPEG	eng2021062301.jpg

- the file metadata Last Modified must be properly filled-in.
- It worth to mention that, even if the filename contains personal data like name and surname (and other personal data may be contained within the file itself, it's important to consider that the anonymisation framework that will be described in the upcoming deliverable D5.1, will ensure that no personal data disclosure will happen.

3.3.2 COMPLEX OBJECT-LIKE DATA

The object like data must be formatted either in JSON or XML format.

Due to the unique features of each CSD, it's necessary that a validation schema is agreed for object-like data. Schemas used to validate the incoming data must be:

- in XSD format for XML serialised data
- in JSON Schema format for JSON serialised data

Whichever new and unforeseen RWD types and/or further elaboration needs that may arise during the project execution will be managed after specific agreements among WP3 and WP5 partners.

The new specifications will be documented in the *D3.2 - Project Manual – CoBraD RWD, Updated MES-CoBraD Protocols and Quality and Quantity (Q&Q) Evaluation – v2*.

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

3.4 Data Entry Considerations for RWD Acquired by a Clnician

There are **two User-Interface (UI) methods** for acquiring RWD:

1. Direct-to-Platform (DtP)
2. Paper-to-Platform (PtP)

It is well established that there are fewer mistakes and omissions through DtP RWD acquisition, such as limiting answers to acceptable values or mandating answers be completed to proceed. Nonetheless, in many settings this is not feasible, especially when considering vulnerable populations with little electronic education and experience, when resources to electronic acquisition tools (computers, tablets, smart-phones etc.) are limited, or, even, when electronic acquisition tools fail (e.g., programs crashing, drained batteries). In such cases, PtP acquisitions are required as alternate or backup methods.

Although either method can bypass the mediation of clinical-research staff during RWD acquisition, in many cases staff are required to clarify questions, especially as novel protocols are being developed and streamlined. This may be achieved asynchronously, however, after patients and caregivers have provided the respective RWD, and following Protocol guidelines to optimise staff time and cost.

In the case of PtP acquisitions, selection of paper-based questions according to an Expert System (ES) - derived decision tree may not be feasible, in which case instructions in paper-based questions will direct to the next set of questions that should be answered.

Irrespective of PtP or DtP UI methods, MES-CoBraD questions can be structured with internal dependencies that allow selection of subsequent questions based on answers to superordinate questions in a decision tree. The details of these dependencies for each question are documented in the dynamically evolving RWD Appendix throughout the project's lifetime, and allow the flexible integration of more complex chronic conditions (CCC) on the MES-CoBraD Platform in the future. See also 4.1.3below.

Finally, for new and ongoing protocols, most RWD will be accessed directly through the MES-CoBraD Platform according to harmonised protocols, however, historical RWD will be accessed **through existing databases**, which usually have a different structure between organisations (private or public), and require implementing harmonisers developed through the respective engineering packages according to platform interoperability protocols (see WP5). Generally speaking, RWD database extraction has two categories of data for extraction: (1) structured variables in a multidimensional tabular form, and (2) natural language text-derived RWD. The former is the case for most databases that have RWD structured in tabular form, and in which acceptable response parameters and quality of responses may be available as part of the respective database documentation and interoperability protocol. Text-derived RWD do not yet have standard protocols in deriving pertinent clinical information as represented in a clinician's note, yet the MES-CoBraD Advanced Analytics modules (WP6) could be leveraged to extract variables from text and harmonising them according to MES-CoBraD RWD variable features.

For all variables, the following parameters are defined, that can also be used in Platform variable developer tools.

1. Name (Cat_Var_User_Input_version) (required)
    a. Cat reflects the category of RWD the variable represents (e.g., CSH for Clinical Symptoms and History)
    b. Var reflects the core information of the variable (e.g., Misplace for misplacing objects)
    c. User is usually patient, caregiver, or clinician-scientist that provides the information, not the mediator that enters it (e.g., if the question is directed to the patient, it will be Patient, even if the clinician-scientist enters it)

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

   d. Input is the method of acquiring the information (Q for question to user, PE for physical exam by clinician-scientist, NPT for neuropsychological testing by neuropsychologist etc.)

   e. Version is the version of the variable as it may differ in one or more of its following parameters, despite retaining the same core semantic information

2. Description (required)
   a. Free text describing what the variable represents for clinician-scientists to understand what the variable represents
3. User_interface (required for question-trees and questionnaires)
   a. Text presented to the user in either electronic (DtP) or paper (PtP) format to input the value of the variable
4. Type (required)
   a. Type of data: integer, real, Boolean, text
5. Measure (required)
   a. Continuous, ordinal, categorical
6. Max / min (required if numeric or if acceptable_values not completed)
   a. Maximum and minimum acceptable value
7. Acceptable_values (required if max /min not completed)
   a. Specific acceptable values for variable. Usually for categorical or ordinal data (e.g., “Yes/No”, or “Never, Rarely, Often, Always”)
8. Missing_value_code (required)
   a. Empty or -99
9. No_response_code (required)
   a. Empty or -88. Indicative that a person chose not to provide data for a specific variables, potentially useful for certain variables
10. Grouping (at least one required)
   a. One or more categorical groupings based on Data Abstraction hierarchical trees (see Appendix 1). Useful in pursuing analyses afterwards based on semantic associations between variables
11. Questionnaire_assignment (optional)
   a. For variables having question versions belonging to one or more questionnaires or tests. Useful for extracting information that can be harmonised between databases, as well as minimising repetitiveness of questions and tests within a project.
12. Conditional_dependencies (optional)
   a. Criterion for questions belonging in a question tree to be asked if one or more superordinate questions/variables satisfy a condition (e.g., if “CSH_SeizureHistory_Patient_Question_v1 == Yes” indicating that a person reported having seizures in the past, then User_interface in variable “CSH_SeizureFrequency_Patient_Question_v1” will be displayed for variable

The MES-CoBraD project has received funding from the European Union’s Horizon 2020 Research and Innovation Programme under grant agreement No 965422

completion of how often they have seizures)

b. For multiple dependencies, AND and OR operands can link clauses, and nested with parentheses

Note that variables that are composites of other variables (e.g., ISI total score for insomnia), this information is derived within the platform as a first-order latent variable and is not part of the input to the Platform by users to minimise data entry errors. Instead, individual zero-order latent variables that are required to be entered by users are the ones represented in the Appendix.

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

4 REAL WORLD DATA CATEGORIES

4.1 CLINIC SYMPTOMS AND HISTORY

Clinical Symptoms and History (CSH) RWD are the cornerstone RWD of any assessment, and represent individuals' salient complaints related to CoBraD, e.g memory problems in the setting of NCD, as well as historical clinical and social information that contribute to or are affected by symptoms. Symptom relief is also one of the most important benchmarks for a successful therapy. Due to the nature of CoBraD, multiple symptoms may occur in a single person, may manifest with variable severity, and differentially impact quality of life. Moreover, symptoms may be caused by other comorbid medical conditions or their therapies. The high level of complexity and dimensionality of CSH requires systematic deep phenotyping of symptoms and history, ideally by trained experts as per current state-of-the-art, and subsequently integrated with findings on physical exam and laboratory testing. The CSH can be provided directly by patients or their caregivers, or indirectly by clinical-research staff, and represent:

1. Troubling symptoms of a patient with regards to their History of Present Illness (HPI) following a Chief Complaint (CC), current Review of Systems (ROS), and current Functional Status (FS), as well as
2. Historical information of Past Medical and Developmental History (PMH), Social History (SH), Family History (FH), and Medications

This type of information allows clinicians and researchers to identify features in a person's symptoms and history that allow deep-phenotyping of a patient's medical and social condition, including a possible underlying pathophysiological process, and guide further workup towards accurate diagnoses and treatment choices.

4.1.1 SOURCE, ACQUISITION METHOD, DATA ABSTRACTION

There are **four main sources** from where CSH are derived:

1. Patients
2. Caregivers
3. Clinicians-researchers and clinical-research coordinators
4. Existing databases

Most of the CSH RWD are acquired directly from patients or their caregivers, whereas certain information reflect metrics entered by clinician-scientists or clinical-research coordinators after interviewing patient-caregiver dyads. CSH RWD only reflect patient-caregiver perceptions and not clinician-scientist or coordinator assessments during an encounter with patients-caregivers, which are represented through Physical Examination & Neuropsychological Testing RWD. Under this premise, clinical-scientist staff serve as facilitators and, accordingly, mediators in CSH RWD acquisition.

There are **three acquisition formats** for CSH RWD:

1. Structured questionnaires
2. Decision tree specialty-specific questions
3. Database extraction (e.g., Electronic Medical Records [EMR])

All formats can represent the same underlying information, so that a question in a decision tree can also be a question of a questionnaire, or be represented as a variable in a database. The high-level RWD semantic associations that MES-CoBraD is built to account for allows Advanced Analytics modules easier data integration and analysis, that are subsequently exploited through its Expert System.

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Structured questionnaires are series of predefined questions provided in sequence to a patient or their caregiver to answer according to specific instructions. In many cases, a questionnaire score is computed through a weighted sum of all questionnaire answers, representing CSH severity at a given time, and, thus, facilitating follow up of CSH over time. This also explains why questionnaires are almost always used in clinical research and trials as primary or secondary endpoints. The MES-CoBraD Platform allows a clinician-scientist to choose through their protocols one or more structured questionnaires, or have derived questionnaire scores from answers provided through specialty-specific question trees or database information as described below. Most questionnaires represent in one form or another the information obtained through ROS, FS, PMH, FH, SH, and Medications.

When it comes to the reason a person visited the clinic in the first place, however, exhaustively going through arbitrary questionnaires is not practical, and can be misleading. Instead, when a patient visits their clinician, they provide CSH RWD according to **question trees**, where the answer of one question leads to deciding which is the best next question. This is best represented in a clinician's HPI questions following a person's answer to a Chief Complaint (CC) for visiting a clinic. These series of questions can be identical to the individual elements of one or more questionnaires, but they are asked (a) only if pertinent according to previous questions, and (b) often relate more to the CC. Nonetheless, if a patient or their caregiver responds to question trees that include all questions of a questionnaire, an estimated score of that questionnaire can also be derived.

Note that different CSH RWD acquisition formats can lead to different responses by users, however, as long as the question is understood by the user (patient, caregiver, or clinician-scientist), these differences are small in most cases.

Data abstraction of CSH RWD is based on three orthogonal categories, each having a hierarchical structure:

1. CSH-based RWD
2. STS (Severity, Temporal, Spatial)
3. CCC-specific RWD (i.e., CoBraD-specific RWD)

The orthogonal nature of the above categories allows for RWD to be classified on several dimensions, thus facilitating the grouping of variables and identifying semantic associations between them for future analyses by clinician-scientists and Expert Systems.

All three data abstraction categories and their hierarchical structure are expanded on in ANNEX I :

The **CSH-based RWD** category structure is fixed for all types of CCC, including CoBraD, and reflects principles of the clinician's interview, organised as subcategories.

Such a categorisation allows for deciding on the sequence and grouping of presenting CSH RWD to users, verifying that CSH-based RWD are presented and completed by a user, thus optimising completeness of data acquisition in clinic or research.

The **STS** category, in a self-explanatory manner, indicates if CSH RWD information refers to presence/absence and severity of CSH RWD (i.e., Severity), whether it refers to the duration, periodicity, and speed of symptom progression (i.e., Temporal), or the spatial features of these symptoms (i.e., Spatial). Note that presence of RWD that reflect Temporal and Spatial features are conditional on presence of RWD Severity.

The **CCC-specific** RWD category data abstraction is variable according to the specific CCC a project aims to address. For the MES-CoBraD pilot, hierarchical data abstraction of CoBraD represents CSH relevant to neurocognitive, epilepsy, and sleep disorders. In line with the conceptual framework of the MES-CoBraD Project, the CCC-specific category domains (i.e., hierarchal data abstraction) are modular in nature, allowing tailoring to varied clinical-research settings and new evolving scientific knowledge. The

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

level of depth within this hierarchical structure between superordinate to subordinate levels is similarly modifiable and varies according to the needs of the clinicians and scientists. Lower levels of abstraction (subordinate) allow for finer-grained information description of CSH RWD.

Note that CCC-specific category organisation corresponds to CCC-specific hierarchical data abstraction of clinical examinations and laboratory tests applied in clinical-research practice to objectively examine these CCC-specific categories. For example, neuropsychological tests reflect metrics of cognition, whereas sleep studies metrics of sleep. To that end, there is a large overlap of the respective data abstraction between categories of RWD (see respective chapters of the MES-CoBraD Manual).

For example, a CSH variable relating to "seizure frequency" can be grouped in three dimensions as HPI (CSH-based RWD), Severity, and Frequency and Rhythmicity (STS), and Epilepsy > Seizure (CCC-specific RWD).

4.1.2 QUALITY CONTROL

Quality control of CSH RWD reflects

1. Responses within a prespecified value range
2. Completeness of data within protocol requirements
3. Correspondence of user responses to the intended construct

Responses within a prespecified value range are easily controlled for DtP acquisitions where the acceptable range of responses is provided in the Appendix table with the respective variable's elements.

Instead, errors are more likely to occur in PtP acquisitions, where participants may (a) circle multiple responses on a sheet of paper (e.g., "4" and "5" when they feel a value of "4.5" is more representative; or "4," "5," and "6" when they experience more variability in their symptom perception), or respond with value types different from the question at hand (e.g., "a lot" as an answer to "What year did your symptoms start?") often reflecting poor comprehension. These types of errors will be mitigated by having clinical or research coordinators review together with participants their answers, targeting specific CSH RWD where paper responses are invalid and providing clarifications. If an invalid response is missed, then the acceptable response will be the median or mean value (depending on data type). If this is not possible, then the response is considered invalid, and the respective code value (e.g., -99) can be entered instead. Although most CSH RWD variables have the same missing or erroneous value codes (e.g., -99 or -88), these are provided for each variable within the respective Appendix.

Completeness of data within protocol requirements

Clinical or research protocols require that participants provide a set of information to allow diagnosis or direct treatment decisions. Completeness of data falls under a protocol's RWD quantity requirements and as RWD is entered on the Platform, coordination tools can inform of missing information to be acquired.

Additionally, unique to CSH RWD (as well as certain clinical and neuropsychological examination RWD), information is acquired through a sequence of questions when the acquisition format is a questionnaire or a question tree (see above). In the case of DtP acquisition, the Platform itself is able to direct the users to answer unanswered mandatory questions before the protocol proceeds with the next question. In the case of PtP acquisition, a clinical-research coordinator or a clinician will have to ask the user (usually a patient or a caregiver) to complete the unanswered questions.

Correspondence of user responses to the intended construct

In addition to quality control of RWD input, an often-overlooked component of quality control is whether CSH RWD represent the intended construct, i.e., information. To mitigate this, all questions through which CSH RWD are acquired are vetted through pilot organisations, and translated and localised to their regional populations and assessed during pilot protocol test runs. Furthermore, participants will be

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

regularly assessed through clinicians on their responses, and over time (either during protocol test runs or pilot period) poor construct representations will be flagged and corrected. Finally, construct validity of individual or composite variables can also be pursued through Platform-based analyses.

4.1.3 Data Storage

CSH RWD entered directly in an anonymised manner DtP will be available to the Platform, and at the same time will be available to the local site's secure and GDPR compliant repository. CSH RWD entered via PtD, will have paper-based responses scanned in pdf format on the local site's secure and GDPR compliant repository, and clinical-research coordinators will enter the information in an asynchronous manner to the Platform.

4.2 Physical Examination and Neuropsychological Testing

Physical Examination (PE) and Neuropsychological Testing (NPT) contain information obtained by trained examiners on physical signs and mental abilities. In contrast to CSH, where the information is obtained from patients or caregivers, the information in PE and NPT is obtained only from clinician-scientists based on their expertise to assess patients. The modalities of assessment fall under the CCC-based RWD category as reported in Appendix 1 and correspond to CSH information, allowing for an objective assessment of reported symptoms and future analyses between types of RWD.

PE within MES-CoBraD is guided by relevant signs that relate to specific CoBraD (e.g., bradykinesia, rigidity, postural instability, tremor as features of synuclein-related neurocognitive disorders, or enlarged tonsils for narrow upper airway as a contributor to sleep apnea). For purposes of harmonisation and diagnostic utility, metrics of PE are tailored to tests helpful in differential diagnosis within CoBraD, as well as those that allow composite scores to assess disease progression.

NPT are used to objectively evaluate a person's cognitive functions (usually divided into memory, executive, visuospatial, and language). Such tests provide accurate and specific information in relation cognitive abilities, which guides diagnosis and, occasionally, prognosis. NPT are administered to patients with CoBraD by staff trained in neuropsychological assessments and represent:

1. age and education appropriate NPT, whose performance is usually summarised in a single raw variable or standardised score (Z-score), or
2. qualitative data, which are not typically captured in a specific score, and may relay additional information. E.g., the sequence of recalled words from a list, or the type of errors made in copying a shape. Such qualitative data could be quantified, but are not yet for many tests, and are instead provided either as categorical variables or unstructured free text.

The above information allows clinicians and researchers to identify impaired domains that may correlate to a person's CSH features and allow localising changes to certain networks of the nervous system.

4.2.1 Source, Acquisition Method, Data Abstraction

PE is to be performed directly on a patient by a board-certified clinician following the basics of the Neurological examination, covering all relevant signs within the abstracted subcategories of CCC-based RWD (Appendix 1).

Similarly, NPT are to be administered through neuropsychological batteries (i.e., sets of tests), since it is important to identify useful tests, as well as their sequence in administration, especially for tests requiring interference tasks between their steps (e.g., short-term memory tests requiring working memory interference tests of the same modality between registration and recall).

Within the scope of general harmonisation and in keeping with the philosophy of the MES-CoBraD Project

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

of being dynamic and tailored to population needs, (a) the majority of NPT administered will be identical (~70%) and capturing all main cognitive domains across sites after being localised to different populations, (b) a subset of tests will be quasi-identical (e.g., Trails B and modified Trails B), and (c) some will represent the same high-order semantic dimension but through different tests (e.g., two different tests assessing a specific cognitive function). This also facilitates within-site continuity and congruence of RWD acquisition, and provides a research substrate to assess the level of non-identity that is allowable in large multi-site studies. There are several statistical techniques to allow comparability between tests either at the low-level of quasi-identical NPT or higher-level of cognitive domains (e.g., via z-scoring and dimensionality reduction) that are under the purview of WP4 and WP6.

Note: The same philosophy in developing and achieving harmonisation in real-world practice across sites applies for all RWD types.

4.2.1.1 Source and Acquisition method

PE and NPT will be performed during in-person visits with patients. NPT will be performed through paper-and-pencil or electronic-based tests (e.g., tablet-based) for the purposes of the main Pilot Use Cases, although certain sub-projects may allow for caregiver assessments following informed consent that is in line with Institutional Review Board approval documents. Trained assessors in the specific batteries of NPT will be conducting the testing at each site, even if they are not board-certified neuropsychologists, but fulfill the Project's quality control protocols. In contrast, PE will be performed by board-certified clinicians. PE and NPT values will be documented either DtP or through PtP transfer.

Note: As with other RWD, data can be transferred to the MES-CoBraD Platform either by the user directly (clinician-scientist in this case), or be transferred in aggregate via a site's existing database Platform (see Chapter 3 above).

4.2.1.2 Data Abstraction

Data abstraction of PE and NPT follow the same principles as CSH RWD of a hierarchical structure within four orthogonal categories:

1. H&P-based RWD
2. STS (Severity, Temporal, Spatial)
3. CCC-specific RWD (i.e., CoBraD-specific RWD)
4. Data Quality

In addition to the first three categories of data abstraction, PE and NPT allow for direct quality assessment of RWD as they were acquired, reflecting parameters such as degree of cooperation and effort-independent procedure quality. In contrast to CSH RWD, PE and NPT category for STS only reflects Severity, given both types of RWD represent a snapshot of a person's performance.

For example, an NPT variable, such as Mini Mental state exam (MMSE), can be grouped in four dimensions representing neuropsychological examination (H&P-based RWD), Severity (STS), global cognition (CCC-specific RWD), and quality of data acquisition. Similarly, a person who is providing little effort on PE Strength testing allows for their PE to be interpreted within context of poor quality.

4.2.2 Quality Control

The following guidelines are provided to maintain inter-rater reliability and ensure standard administration of PE and NPT. Following these guidelines is helpful to generate valid and accurate measurements with a minimum of stress and discomfort to participants.

Preparation

1. Examiner can influence testing to some degree even when standardised procedures are

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

used, so it is desirable to have the same examiner conduct each of PE and NPT.

2. It is important that the testing takes place in a quiet room free of distractions and with the necessary setup (e.g., examination bed, desk).
3. Before testing, both the participant and their caregiver are to be asked about the participant's ability to hear and see in order to understand commands and perform tasks independent of primary sensory deficits. It is the examiner's responsibility to see that the participant understands the instructions before each test is started and that this understanding is maintained throughout the test. It is, thus, important that the participant is wearing corrective eyeglasses or hearing aids as needed.
4. Both the PE and NPT battery are to be administered for each person in the same order and with adherence to time limits and standardised instructions.
5. Test Administration
6. Instructions should be clear, and for NPT read verbatim without paraphrasing. It is acceptable for instructions to be paraphrased if verbatim instructions are poorly understood, as long as the basic concept is maintained. Instructions may be repeated according to each task's instructions, avoiding additional information, hints, or answers.
7. Any feedback to the participants will be neutral, avoiding to indicate if their answers are correct or not, or whether they are performing within or below normative limits. To that extent, it is prudent to explain to all examinees that feedback will be given at the end of a complete assessment, setting expectations and maximising attention during tasks.
8. If a participant provides more than one response to a NPT, the examiner will encourage them to choose one without cueing for a specific response.
9. It is important to maximise effort throughout an examination. To that end
    a. Examiners should provide regular encouragement as people advance during their assessment to maintain participants' effort. Participants who demonstrate poor effort should have their data categorised as of poor quality.
    b. If participants give up easily or respond with generic avoidant answers (e.g., "I don't know") examiners will try and elicit answers by providing encouraging statements (e.g., "Just guess" or "Give yourself a minute to think of an answer").
    c. Validating frustration can be helpful and build rapport (e.g., "It is anticipated to miss some," and "Even if you miss a few things, that information may help us figure out what your strengths and weaknesses are.")
10. Examiners should aim for participants to complete all assessments even if participants are uncertain of a response. It is more informative to have a wrong answer than no answer
11. An examiner can use personal notes when uncertain how to score a response, and review the PM later or discuss it offline with partners on how to best document information
12. Qualitative observations are useful in identifying nuances between identical quantitative responses, especially when a nuanced metric is not part of a protocol. To that end, text notes can be useful in including or even developing additional variables for assessment of participants (e.g., impulsivity, stimulus-boundness, slowed processing speed, tremor, apraxia).

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

4.2.3 DATA QUALITY CONTROL

1. It is the responsibility of each clinician-scientist conducting an assessment to guarantee the quality and integrity of the data beyond quality control measurements post data entry to the Platform.
2. As with other types of RWD, PE and NPT acceptable values and ranges are entered for each variable on the Platform for real-time quality control upon data entry to the Platform.

4.2.4 DATA STORAGE

Data will be uploaded securely post-anonymisation to the MES-CoBraD Platform either through DtP or PtP processes. File formats are subject to Data structure guidelines (see Chapter 3 of the PM).

Data will be digitally stored for a minimum of 5 years at each primary site and will also be available on the MES-CoBraD Platform for the duration of the Platform's lifetime. Physical records obtained during the examination will be also stored for 5 years. If physical records are scanned on-site for easy access, the paper copy can be discarded as per GDPR guidelines.

4.3 NEUROIMAGING

Neuroimaging is the discipline that deals with the in vivo depiction of anatomy and function of the central nervous system (CNS) in health and disease. There are several methods to image the brain, the most used are brain computed tomography (CT) and Magnetic response imaging (MRI). Brain CT, is the most readily available and cheapest neuroimaging modality and thus typically the first choice in most clinical settings. However, CT scans have low capability to separate brain white and grey matter, to properly annotate brain regions of interest (ROI), and, thus, produce scans with low quality and many imaging artifacts, while also introducing ionizing radiation to participants. For these reasons, CT scan are traditionally not employed in clinical or research settings other than in the acute management of patients, and instead most clinical diagnostic guidelines recommend MRI as the neuroimaging modality of choice. Furthermore, it is the method of choice for quantitively assessing brain structure and function in research settings. In certain settings, primarily for research purposes, Positron Emission Tomography (PET) is a neuroimaging technique often utilised to identify either specific pathology in the brain, especially proteinopathy, or functional activity between brain regions.

4.3.1 SOURCE, ACQUISITION METHOD AND DATA ABSTRACTION

MRI images will be acquired in 1.5 (or higher) Tesla scanners across sites. Each site should have designated responsible personnel that oversee neuroimaging processes, aiming to prevent and manage adverse events.

In what follows, variables are grouped according to low-order variables used in common practice vs. advanced research settings. Higher-order variables obtained post-processing of lower-order data through advanced analytics modules do not fall under the PM purview.

4.3.1.1 Low-order variables: Common

1. T1-based MRI. 3D (Axial, Sagittal, Coronal). At most 1.5 x 1.5 x 1.5 mm voxel size
2. Diffusion weighted Imaging (DWI). Axial. At least one b=0. At least 16 gradient maps (32 desirable) at b=1000, sampling the whole sphere
3. T2-based MRI. Axial and Coronal.
4. FLAIR 3D (Axial, Sagittal, Coronal).

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

5. Gradient Echo or Susceptibility Weighted Imaging. Axial. To assess for (micro)hemorrhages

Common clinical-based acquisition: Research based neuroimaging acquisition are disadvantageous to clinical scans as they do not include acquisition with contrast. Moreover, research neuroimaging studies are traditionally not evaluated by an expert neuroradiologist in a timely manner. For this reason, every pilot partner site should have a minimum, time efficient, non-research-based neuroimaging protocol that allows fast diagnosis of urgent findings (e.g., hemorrhage or brain tumor).

4.3.1.2 Low-order variables: Advanced

1. Diffusion weighted Imaging. 64+ b1000 gradient directions for tractography. Multi-shell for multi-compartment modeling. Including one b=0 with inverse encoding to correct for eddy-current/movement artifacts
2. BOLD. Resting-state functional MRI. 80+ volumes. Eyes closed during acquisition.
3. Additional susceptibility sequences. Ideal to assess microstructure alterations. More than 4 diffusion schemes. Not very robust to parameter acquisition, and post-processing harmonisation is needed.

4.3.1.3 Positron Emission Tomography (PET)

1. Fluorodeoxyglucose (FDG) PET is a versatile test that has been widely used in Alzheimer Disease, Epilepsy, and Sleep disorders.
2. Amyloid and tau PET can also be performed through 2nd generation tracers when considering neurodegenerative processes.
3. DATSCAN can be pursued for people with parkinsonism or suspected synucleinopathy

4.3.2 Data Abstraction

Considering that Neuroimaging data represent CSD, their abstraction primarily reflects a hierarchical categorisation according to the main imaging modality (e.g., MRI) and its subordinate techniques/sequences applied (e.g., T1, Axial), as well as categorising to the quality of the acquired data (see 4.3.3 as well).

MRI and PET images will be exported from the different scanners to a local repository using the within-machine exporting software. Exported data (both MRI and PET) will be stored using the DICOM format.

4.3.3 Quality Control

A stage process will be performed to guarantee the quality of the data. Importantly, such process will report on raw data quality, not pre-processing downstream steps.

4.3.3.1 In-site QC

A trained neuro-radiologist, or a neurologist with MRI expertise should visually inspect the acquired data to guarantee that there are not major movement or other artifacts. If artifacts are present, these will be assessed by the reader and the quality of a sequence will be quantified.

4.3.3.2 Platform QC

There are several automated tools to perform automatic QC assessment that will be integrated on the Platform. One of the most widely used is MRIQC (for T1, T2 and BOLD images) (https://mriqc.readthedocs.io/en/latest/index.html). The computed outcomes from the QC processing can be used to establish thresholds for usability, or to train machine learning models to automatically

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

prompt a quality status of the input images. Analytics modules of movement correction can be applied after images are uploaded on the Platform.

Anonymisation and defacing

Considering the advancement of facial reconstruction from MRI images, to ascertain anonymisation of participant imaging data, modules of defacing and anonymising DICOM images will be implemented prior to uploading images on the Platform.

4.3.4 DATA STORAGE

4.3.4.1 Site Local Storage

Each site is expected to store its acquired data following its own protocols. It is, nonetheless, recommended to use a predefined data structure such as BIDS (https://bids.neuroimaging.io/). Data must be stored in a format closest to low-order data as possible (i.e., DICOM format). In addition, this low-level of data can be converted to the conventional NIfTI format (https://nifti.nimh.nih.gov/). If an imaging acquisition center does not provide backup, each pilot partner is responsible to have a backup process.

4.3.4.2 Cloud-based Platform Storage

Within the Platform, data will be organised according to open source XNAT (www.xnat.org), which is widely used in the neuroimaging community due to its potential to organise both low-order and higher-order (processed) data, in addition to including metadata such as quality control outputs. XNAT is also able to support anonymised data (doi: 10.1016/j.neuroimage.2021.117845).

4.4 BIOLOGICAL SAMPLES

To evaluate and understand how complex brain disorders impact the normal function of the central nervous system and peripheral organs and tissues, cerebrospinal fluid and blood samples from clinical cases and matched controls must be collected, processed, aliquoted, and appropriately stored, in accordance with a harmonised protocol across different centers involved in this consortium. This process allowes determining significant differences within and between groups for the specified parameters and interventions, based on state-of-the-art molecular and biochemical analyses.

4.4.1 BLOOD AND CEREBROSPINAL FLUID

4.4.1.1 Source, Acquisition, Data abstraction

CSF and blood biosamples will be collected from participants by a trained and competent staff member who follows the respective sampling protocol.

Before drawing samples, the staff will ensure that all collection and aliquot tubes are marked with a water- and ethanol-resistant marker and that all text is easily readable. Collection/aliquot tubes should include the following labeling:

1. Participant ID.
2. Intervention ID (e.g., experiment and site).
3. Sample ID (e.g., session, timepoint).
4. Aliquot ID (e.g., serum, plasma, or CSF; also adding aliquot identification number - 1 through 5).

Example:

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

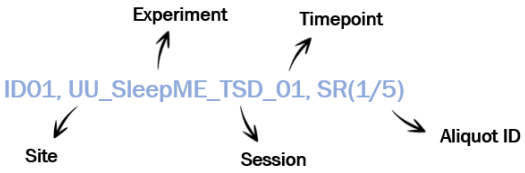


Figure 3 Proper annotation of bio-sample.

Cerebrospinal fluid sampling: A lumbar puncture (LP) is preferably performed in the morning and the subject does not need to be in a fasting state. An LP is performed between L3-L4 or L4-L5 spinal vertebra interspaces, the patient must be sitting or lying down. One catheter/tube can be connected to the needle extractive CSF and sample is then acquired in a 10 mL polypropylene sterile tube.

In general, the few first drops of CSF are discarded to minimise possible contamination. In the case of apparent (visible) blood contamination of the CSF, a larger portion is discarded, and the CSF to be sampled is collected in a new tube (i.e., not in the initial "waste" tube), when the bleeding has stopped. After collection, CSF must be homogenised by turning the tube over several times. Subsequently, the CSF sample is centrifuged at 2200 relative centrifugal force (rcf G-Force) for 10 minutes with temperature set at 20C (or in a similar normal room temperature range).

Blood sampling: when handling blood samples, staff should consider wearing and eye shield, face mask, gloves, and maintain good hand hygiene. Blood should always be considered contagious, and specific modified protocols may apply at each site for samples obtained as a result of the COVID-19 pandemic. In the MES-CoBraD Protocol, we refer to the most common colors of blood tube caps used in practice.

Note that lid colors and tube characteristics (i.e., presence or absence of certain compounds like lithium heparin, sodium heparin, silica, and others) can vary depending on the manufacturer, and some samples can be taken in several types of tubes. It is crucial to verify the correct collection tube type before starting sample collection.

1. It is recommended to use gel separator tubes, also known **as Serum Separator Tube** or *SST tubes* (yellow cap) and **Plasma Separator Tube** or *PST* tubes (light-green cap), since they provide a simple and efficient way to separate the liquid portion of the blood (i.e serum and plasma,) with minimal risk of collecting erythrocytes or their products due to improper pipetting. However, the use of SST and PST precludes proteomics analysis, but for most purposes, this is an acceptable trade-off.
2. **EDTA** (purple cap) and **PAXgene** (clear-brown cap) tubes are ideal for DNA and RNA analysis, respectively.

When starting drawing blood, it is expected that the first portion drawn contains hemolyzed blood, so it is recommended to discard the first 1-2 mL of blood in a waste tube (i.e., regular tube, non-sterile), before switching to a collection tube. Also, to avoid hemolysis, pressure should not be placed around the arm when sampling. For a filled collection tube of 5 mL, 2 mL of serum or plasma can be acquired, allowing approximately six aliquots of 250 µl each.

Post-collection processing

After collection, specific protocols need to be followed depending on the type of sample/tubes.

1. PST can be directly centrifuged, whereas SST must be held in an upright (or semi-upright) position at room temperature for a minimum of 30-35 min (maximum limit is 2 hours) before centrifuging and aliquoting steps.
2. EDTA tubes need to be mixed by turning over several times (8x) before centrifugation. For

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

extracting DNA from whole blood in EDTA tubes, samples must be freezed immediately after mixing them WITHOUT centrifuging. For this purpose, each tube must be dipped into a freezing solution ("slurry", i.e., dry ice + 70% ethanol; the solution is at ~-80°C once dry ice stops "boiling off" from this mix).

3. PAXgene tubes must NOT be centrifuged in most cases. These tubes are to be kept first for 2 hours at room temperature and after that to be placed in a -20°C freezer (for at least 24 hrs), and then for long term storage at -80°C.
4. All centrifugation steps need to be done at 2000 relative centrifugal force (rcf; G-Force) for 10 minutes with temperature set at 20°C.

Note: During sample collection, always document the following:

1. Volume pipetted per aliquot (e.g. 250 µl)
2. Possible blood contamination (clots or hemolysis)
3. That both lid and side of tube are congruently labeled, since mislabeled lid and/or side of tube: will require re-labeling prior to shipping
4. Possible deviations in processing time (time at room temperature / if not frozen soon after aliquoting)

Aliquoting samples

After centrifugation, pipet the samples into mini tubes (i.e., polypropylene sterile tubes, with total volume of 1.5 ml).

1. Plasma
    a. Pipetting to 6 mini tubes (label them as PR[1-6], see above for full naming convention).
    b. 250 microliters in every mini tube, if not specified otherwise.
2. Serum
    a. Pipetting to 6 mini tubes (label them as SR[1-6], see above for full naming convention).
    b. 250 microliters in every tube, if not specified otherwise.
3. CSF
    a. Pipetting to 6 mini tubes (label them as CSF[1-6], see above for full naming convention).
    b. 500 microliters in every mini tube, if not specified otherwise.

Aliquots must be placed directly into dry ice directly after aliquoting to achieve immediate freezing of the samples, being aware that freezing occurs faster if aliquots are fully submerged. It is, thus, important that there is enough dry ice for the number of aliquots, and for the duration of whole process, since dry ice will evaporate. A few kilograms of dry ice can be stored in styrofoam boxes for some hours at room temperature, which is especially important when collection occurs outside the lab, and for several months in ultrafreezers (-80°C) if they need to be stocked. The latter works best in a lying (chest) freezer. Additionally, minimising the interface between air and dry ice in its container (e.g., by placing newspaper over the dry ice) could further delay its evaporation. When pipetting is completed, aliquots should be transferred to a -80°C freezer as soon as possible.

Once in a -80°C freezer, metabolites and biomolecules in samples usually maintain good integrity for

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

several years. This does, however, require that the freezer does not warm up (briefly recorded higher temperature is fine, e.g., for a few minutes when removing large sets of samples). Temporary freezer warming can be greatly minimised, or entirely avoided, by proper freezer handling technique. Many biomolecules will degrade following even a single freeze-thaw cycle, highlighting the need for proper documentation of sample storage even within the freezer.

As such, clearly label each box you place in the freezer – in a consistent easily identifiable manner – and keep notes in a sampling protocol and/or in a master sample file information about box number/name, site name, freezer number and rack. Clearly label both the bottom and the top (removable lid) of each box.

Do not freeze entire PST or SST tubes: only the final aliquots of serum/plasma, as well as EDTA (for DNA extraction) and PAXgene (for whole-blood RNA extraction) tubes should be frozen.

An overview of the above protocol is illustrated as a flowchart in ANNEX V :.

Shipping samples

Sample analyses for the purposes of MES-CoBraD are performed at Upsaala University (UU), where biomarker analysis from metabolites to genotyping can be pursued. Proteomic analysis (up to 1,536 targets, see ANNEX VI :) can be achieved using a small sample volume, enzyme-linked immunosorbent assays (ELISAs) can be implemented for specific protein targets. More sensitive techniques examining low concentration compounds (femtogram per milliliter), such as single-molecule array technology (SIMOA) are also available. DNA methylation, RNA, and microRNA sequencing are also part of UU partners panel of analysis.

Samples should be shipped to UU on dry ice with the following considerations:

1. Calculate the amount of dry ice depending on the number of days required for shipment (up to 2 days for many parts of Europe)
2. Remember to ship samples carefully labeled, record the whole number of samples and the exact name of each sample/aliquot
3. Ensure that no samples are contaminated by plasma or blood, as they will be discarded
4. Take photos of the sample layout before shipment and attach with the shipment or upload them on the MES-CoBraD Intranet.
5. Include sample description (sample arrangement and description), label any deviations in labeling or content clearly. Also, send an e-mail with samples description and layout to handling partner at UU.
6. Do not allow loose samples to be shipped.
7. Always notify handling partner at UU before shipment and ask for a green light before shipping.
8. The recipient may first need to make room in their freezer, which is also why it is important to condense samples in the shipped boxes, to ensure the minimal amount of space is used.

As samples are collected post COVID-19, they will all be considered potentially hazardous, so detailed information on potential SARS-CoV-2 contamination is also needed (specifying whether the risk is known or unknown). This may also regulate how samples are stored and shipped locally (i.e., if a site is allowed to receive samples). This is an additional reason to ensure that samples are not dirty (e.g., with external blood / serum / plasma contamination), as these may not be handled by the recipient.

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

4.4.1.2 Data abstraction

Data abstraction of biosamples is reflected in the following orthogonal categories:

1. Type of biosample
2. Timing of sample collection
3. Collection Protocol Code

Quality control

It is essential to highlight that laboratory analysis should be repeatable, and if biological specimens are not correctly sampled and/or processed, the final results may be spurious and impossible to replicate. Considering that the MES-CoBraD study will continue for years and recruit a large number of patients, minor deviations from this biosample collection and processing protocol could contribute to substantial data variability and interfere with the study's goals.

For biological specimen collection, the following are important indicators of poor specimen quality, and indicate the need to discard and, if possible, re-acquire the sample:

- Blood contamination in CSF samples
- Notable hemolysis in serum or plasma samples
- Presence of clots in whole blood samples (i.e., EDTA, PAXgene tubes)

Additionally, Inappropriate handling, transport, and storage may be identified post-shipping and analysis by revealing molecular degradation noticeable during biomarker quality check control (i.e., low RNA/DNA integrity).

4.4.1.3 Data storage

Long term storage of biosamples will be pursued in Aliquots, EDTA, or PAXgene tubes within a deep freezer at -80 C, as per aforementioned protocols.

4.4.2 HAIR

Stress is one of the most common seizure triggers reported from patients with epilepsy and there are several experimental and observational data supporting a possible relation between biological stress system, namely Hypothalamus-Pituitary-Adrenal (HPA) axis, dysregulation and epilepsy. HPA axis function is traditionally studied by measurements of its final product, cortisol. Changes in cortisol levels have been linked with all the three CoBraD. There is accumulating evidence that hair cortisol is a valid biomarker of HPA axis function in chronic stress. In contrast to serum and salivary cortisol measurements, cortisol concentrations in hair are not influenced by acute stress, and, instead, assess the integrated levels of cortisol over extended periods of time before sampling. Therefore, hair samples for measurement of cortisol levels will be collected from participants during MES-CoBraD.

4.4.2.1 Source, Acquisition, Data abstraction

Around 150 strands of hair will be collected from the posterior vertex of the scalp, cut off as close to the scalp as possible. The hair will be taped to a piece of paper that will serve also as a document on which all the sample details will be included. The form with the samples will be placed in a paper folder with the identification code of the sample which then will be sealed. The samples can be stored at room temperature, away from ultraviolet (UV) exposure, until they are analysed.

In addition to UU, hair samples can be processed at Choremeion Research Laboratory at the Medical School of University of Athens, Greece, where samples should be shipped. The Proximal 3-cm hair segments from each sample will be weighed (samples approx. between 24 and 28 mg) and placed in

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

grinding tubes, and subsequently lysed at 5000 rpm. Then, the powder-form hair will be extracted in 1 mL methanol at 37 °C while shaking for 16 hours. The extract is transferred to a glass tube and the methanol is evaporated at 52 °C under a constant stream of nitrogen, until the samples are completely dried. Samples are then reconstituted in 100 µL phosphate-buffered saline (pH 8.0, 1 × phosphate buffered saline or PBS) and vortexed for 1 min. Before analysis, samples are vortexed again. Finally, samples are analysed by using automated electrochemiluminescence immunoassay "ECLIA" (see figure below from Wester and Van Rossum 2015). They should be anonymised and coded according to the central code system of MES-CoBraD project and clearly labelled. The samples are not considered hazardous but the sealed envelopes could be placed and sealed in a waterproof box ensuring optimal sample conditions. Given that hair samples do not require analysing in a specific time frame, the samples can be shipped in aggregate every several months.

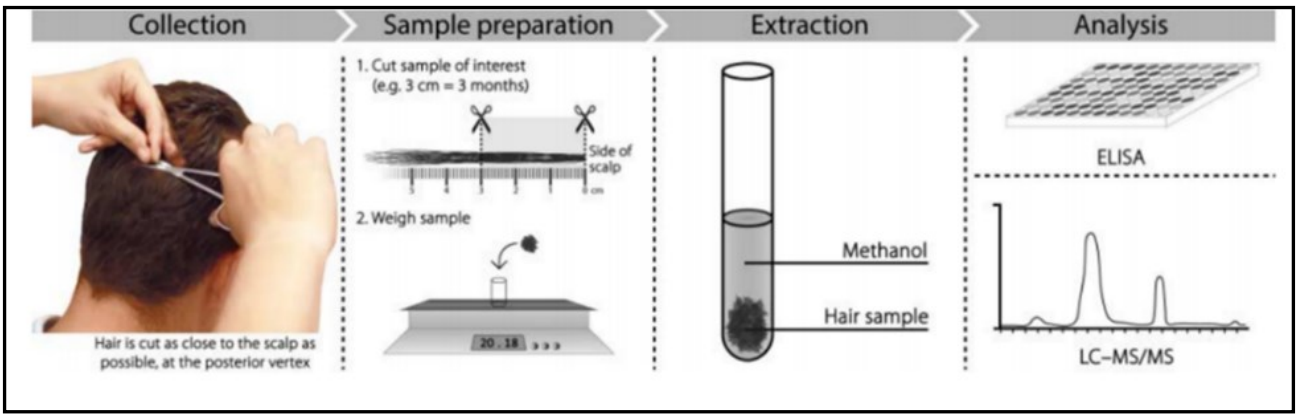


Figure 4 Overview of hair sample collection and analysis

4.4.2.2 Quality Control

The hair should be collected from a clean head on the night of PSG/EEG, and before electrode placement. The hair should not be treated with hair style products during the day of sample collection and if there is hair dye, it should be documented. The side of the hair closest to the scalp should be clearly annotated, because this part will be used for analysis. The envelope containing the hair sample should be sealed and remain intact until reaching the laboratory.

4.4.2.3 Data Storage

Pre-processing, samples are stored as described above until shipping. Processed samples are discarded and hair cortisol results are stored on the MES-CoBraD Platform database. The remaining hair samples are stored for a period of time as per current EU rules.

4.5 PHYSIOLOGY

Physiology measurements within the MES-CoBraD Project represent Electroencephalography (EEG), Polysomnography (PSG), and Actigraphy. Considering the closer association of actigraphy to information acquired through Consumer Technology RWD, it is discussed in the respective chapter (4.7) in more detail. A critical feature of MES-CoBraD is that EEG and PSG are acquired concurrently through a combined protocol. EEG allows for establishment of brain states of alertness and cortical activity, as well as capturing aberrant epileptic activity. PSG protocols exploit EEG information for establishing brain states of alertness, and further supplement through multiple sensors additional aspects of breathing effort and flow, oxygen saturation, snoring, and muscle and eye movement.

4.5.1 SOURCE, ACQUISITION, DATA ABSTRACTION

Source: The data will be obtained from physiological studies performed as part of the multidisciplinary MES-CoBraD Protocol, as well as through pre-existing databases. Data acquired will be exported from the

EEG/PSG database of each partner in anonymised European Data Format (EDF+).

Acquisition: The EEG and sleep recording should have a minimum number and electrodes and positions, see appendix 3. The data will not include patient videos and acquisitions will be anonymised (modified EDF+ header) prior to uploading by utilising Platform tools. Prospective recordings should include a whole night EEG/PSG plus at least one hour of awake resting EEG, minimising artifacts. EEG electrode position in the MES-CoBraD PSG/EEG protocol is based on the standard 10-20 configuration with extra F10, T10, P10 or the Modified Maudsley system (see figure in appendix 3). Electrodes should be placed after scrubbing the scalp and skin with special gel (e.g., NuPrep) and applying EEG paste (e..g, Ten20) while securing electrodes with gauze on which special adhesive paste is applied (e.g., EC2). Expert bandages with paste are also accepted if impedances remain low.

Data abstraction of biosamples is reflected in the following orthogonal categories:

1. Collection Protocol Code
2. Timing of sample collection

As with other types of RWD where variables of practical use are secondary, latent, variables, data abstraction of those variables depends on the post-processing of this information.

4.5.2 QUALITY CONTROL

Quality control will be performed at three different stages. Every time quality control is performed, results are documented to represent success and failure rates and qualitative features for improvement.

Retrospective Physiology Data

1. All studies will be assessed for quality by partners within a site
2. Data will be checked for anonymisation
3. Upon completion of files' quality control, retrospectively obtained and anonymised EDF+ files will be transferred to the Platform.
4. Of the total sample, 5% of data will be sampled for quality control by experts from other partners through visual inspection
5. Reviewers will assess whether recordings satisfy acquisition criteria as outlined above and in appendices
6. In case of systematic errors in a collecting site's database, quality improvements will be recommended by experts from the respective Work Group that reviews physiology data

Prospective Pilot Physiology Data

1. All studies will be assessed for quality by partners within a site
2. Data will be checked for anonymisation
3. Upon completion of files' quality control, retrospectively obtained and anonymised EDF+ files will be transferred to the Platform.
4. Of the total sample, 10% of data will be sampled for quality control by experts from other partners through visual inspection, one reviewer being from KCL for EEG quality control
5. Reviewers will assess whether recordings satisfy acquisition criteria as outlined above and in appendices
6. In case of systematic errors in a collecting site's database, quality improvements will be recommended by experts from the respective Work Group that reviews physiology data

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

4.5.3 Data Storage

Data will be stored post-anonymisation based on unique ID provided for any participant entering the Consortium. EDF+ will be anonymised in a local encrypted storage device for the period specified in each center's guidelines for storage of research data. Files will be stored on the MES-CoBraD Platform for the duration of the Platform's lifetime.

4.6 Medical Device Data

During last four decades a number of medical devices have been licensed for treatment of people with CoBraD who do not respond adequately to drug treatments. Those that aim to address Epilepsy and other nervous system network dysfunctions are often implantable, and their function is exerted through electrical stimulation that modulates signals of the nervous system, collectively termed Neuromodulation. Others address Sleep Disorders, and especially Sleep Disordered Breathing such as Sleep Apnea, usually providing a mechanical intervention through air compression to allow airway patency and facilitate ventilation.

Two main categories of Neuromodulation devices applicable to MES-CoBraD are Vagal Nerve Stimulation (VNS), for the treatment of drug resistant epilepsy and difficult-to-treat depression, and Deep Brain Stimulation (DBS), which is used broadly for movement disorders such as Parkinson disease, essential tremor, and dystonia. DBS is also being assessed in cases of refractory epilepsy, major depressive disorder, and Alzheimer disease. At present, most of these devices do not provide any recording function for continuous data collection. DBS devices with recording capabilities are currently being developed and the MES-CoBraD Project has taken this future potential into account in its planning. In such cases, acquisition of the intracranial DBS with concurrent EEG recordings can also be performed for the purposes of MES-CoBraD. The VNS and DBS variables included in the project can be seen in appendix.

Mechanical breathing interventions that aim to improve upper airway resistance, narrowing, or transient occlusion during sleep are defined as Positive Airway Pressure machines, where an air compressor is placed at the bedstand and through a tubing and mask interface (pillows, nasal, full face), pushes air into a person, thus preventing collapse of the upper airway. In the case where ventilatory support is also to benefit a patient, as may happen in a patient with comorbid lung or neuromuscular disease, then a pressure support is provided by the machine between inspiration and expiration, effectively making it a Non-Invasive Ventilator (NIV) that allows better respiration and ventilation.

4.6.1 Source, Acquisition, Data Abstraction

Patients participating in the MES-CoBraD will be asked of device use and information documented through CSH RWD acquisition. Subsequently, any data that machines provide, especially those concurrent to MES-CoBraD Protocol acquisition, will be acquired following the respective machine's data extraction protocol. See Appendix 3 for details on metrics acquired by machine types. Data that involve clinic assessments and interventions that are not part of machine-provided data (e.g., magnet swaps, side effects) these will be collected through respective question-trees and questionnaires through CSH RWD.

4.6.2 Acquisition and Data Abstraction

Some machines allow direct access to data through APIs, whereas for others there is a need to export data to a digital spreadsheet following RWD structure guidelines (see Chapter 3) prior to uploading to the MES-CoBraD Platform. As per all variables, data will be anonymised and coded according to a central MES-CoBraD code generator and stored in digital format on site prior to uploading on the Platform.

Data abstraction of complex data is represented in Appendix 3 for the main categories of machines. Some machines allow more information than others, possibly allowing for further data abstraction.

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

4.6.3 QUALITY CONTROL

The numerical variables should be expressed in harmonised measurement units as agreed between the respective Work Group partners. Intensity should be expressed in mA, Voltage in mV, pulse width in msec, frequency in Hz. This allows for harmonisation also between machines of a certain type that do not have identical metrics. Randomly chosen cases will be periodically chosen to confirm harmonisation is followed by all partners, although this is not anticipated to be an issue once Platform harmonisation modules are implemented by each site.

4.6.4 DATA STORAGE

Data will be stored locally in encrypted storage media for the period specified in each center's guidelines for storage of research data, and on the MES-CoBraD Platform throughout the Platform's lifetime.

4.7 CONSUMER TECHNOLOGY WEARABLES

Advances in consumer technologies have provided new tools to measure physiological biomarkers such as heart rate captured by plethysmography and activity through accelerometers, from which other secondary metrics are inferred (e.g., sleep and wake states). For example, consumer-based activity trackers (i.e., smartwatches and wrist bands) have increased in popularity in society. Their extended capacity to record data in high resolution makes them a valuable source for epidemiological studies and real-world databases. This chapter will instruct how to acquire, process, and store sleep and activity data from wearables. It is worth noting that the same principles applied in consumer technology wearables for activity monitoring are also applied in clinical-research grade actigraphs, using accelerometers to collect and then transform data directly into a proprietary "activity" metric.

4.7.1 SOURCE, ACQUISITION, DATA ABSTRACTION

Data Source and Acquisition method

Most wearable companies (e.g., Fitbit and Philips) offer different products that can extract high-resolution activity data and secondarily estimate sleep timing and duration. People are required to wear their devices for at least a week to allow a relatively robust pattern of sleep-wake patterns. Most of these devices are also waterproof, with recommendations, however, to avoiding submersion into water for long times. Each device has a proprietary interface for uploading data on either a local application or a cloud-based platform, and some (e.g., Fitbit) allow for API-based data sharing after consent is provided by a user (see Appendix 7). Although many wearables now provide information beyond activity metrics, for the purposes of MES-CoBraD, activity metrics are considered mandatory for acquisition and further analyses, whereas the rest are optional.

4.7.1.1 Acquisition

Data is uploaded to fitbit cloud services. Users may approve access to their data which may be be accessed and download via the internet I (see ANNEX VII :).

After download, raw data (.json files) need to be converted, cleaned, and processed (these steps could be performed using Microsoft Power BI, or algorithms from open-source programming languages like R and Python).

4.7.2 VARIABLES

Physiology data is plotted minute by minute following proprietary algorithms, preventing access to raw sensor data to protect reverse engineering of proprietary algorithms. Any higher-order metrics derived from periodogram analysis and cosine fitting curves like power, mesor, amplitude, and acrophase, provide information about subjects' rest-activity rhythms. Additional variables such as total daily activity,

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

percentages of activity during the day and night, interdaily stability, intradaily variability, mean value of five consecutive hours with lowest activity, mean value of ten consecutive hours with highest activity, relative amplitude, and circadian function index are important variables to complement the characterisation of the rest-activity rhythm in subjects with complex brain disorders. Although these are higher-order variables at face value, they are treated as low-order and primary variables, since the devices do not provide raw data. Especially for sleep metrics, sleep onset, wakeup after sleep onset, time in bed, total sleep duration, sleep midpoint on work and free days, midpoint of sleep on free days corrected for sleep debt on workdays, are important. These variables will allow us to understand the main sleep characteristics affected in most CoBraD.

4.7.3 QUALITY CONTROL

Partners from each study centre must evaluate activity data quality regarding parameters such as data extension (i.e., minimum of one week) and data continuity. Partners are also expected to acquire actigraphy data concurrently with sleep diaries (see Chapter 5), and during the last day of acquisition to have concurrent PSG/EEG.

4.7.4 DATA STORAGE

Data from consumer technologies, multiple variables RWD (MVD), will be stored as coded results in JSON files according to chapter 3, and then direct uploaded to remote portals' APIs.

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

5 REAL WORLD DATA SHARED PROTOCOL FOR THE PILOT DATA

Current state-of-the-art in evaluating people with CoBraD address a patient's specific complaint or syndromic presentation in isolation, not addressing other comorbidities or the multidisciplinary complexity of a syndrome. The lack of comprehensive assessments is one of the most important factors explaining delays in diagnosis as many patients are not aware and clinicians may overlook that a certain symptom is significant to mention in clinical practice, or are simply unaware of concerning signs.

We will pursue the comprehensive assessment of CoBraD through the following novel clinical multidisciplinary multisource diagnostic protocol (Figure 5-1) that relies on optimising real-world practices, by integrating several modalities of RWD and their Metadata, allowing for deep-phenotyping of people with NCD, epilepsy, and sleep disorders, and assessment of their comorbidities, towards achieving precision and personalised medicine goals. The below clinical protocol for multisource multidisciplinary RWD acquisition and integration within and between CoBraD from multiple institutions addresses challenges of CoBraD underdiagnosis, misdiagnosis, suboptimal diagnosis, and delay in diagnosis of primary and comorbid CoBraD, by considering the complex multidimensional CoBraD pathophysiology and exploiting the combined expertise of leading researchers in their respective field from across Europe.

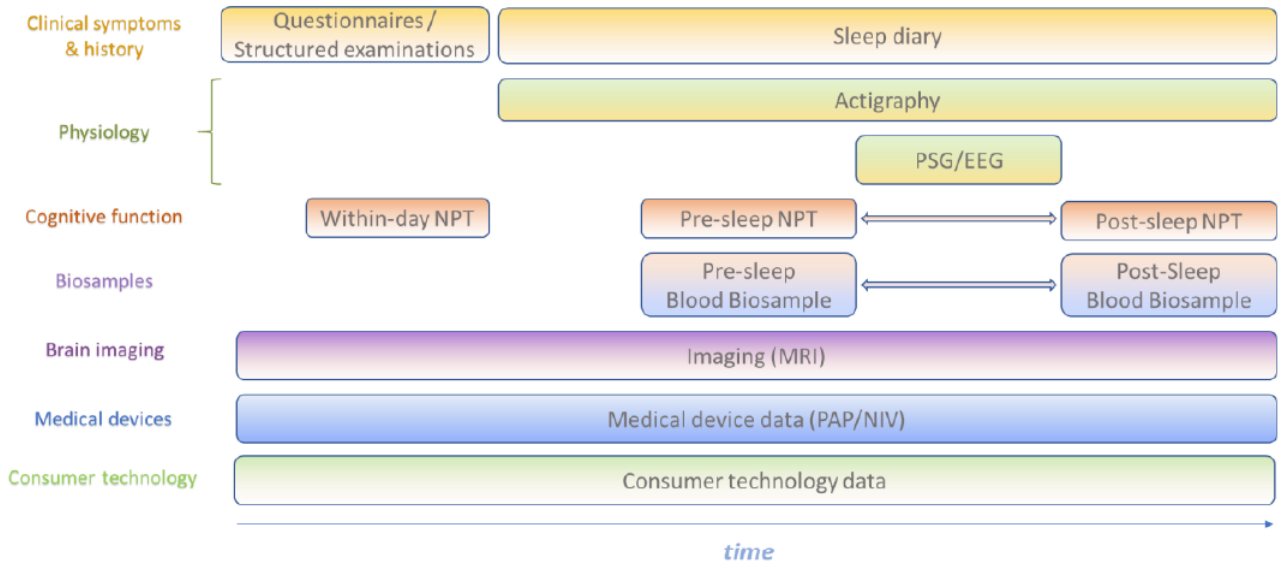


Figure 5 MES-CoBraD Multidisciplinary Multisource RWD Assessment Protocol.

RWD will be acquired from multiple sources and reflective of multiple disciplines of all CoBraD in a temporally structured flow within a week, to maximise their impact in deep-phenotyping of CoBraD. Sleep-mediated evaluations, such as cognitive and blood testing, are significant on their own, but also of their temporal variability throughout the circadian rhythm. Questionnaire data will be obtained in clinic or out-of-clinic by patients, caregivers, and clinicians. Imaging, medical device, and consumer technology data, will be recorded as acquired in real-world settings throughout the project's duration, accounting for timing of their acquisition to other modules. Note: NPT: Neuropsychological Testing; MRI: Magnetic Resonance Imaging; PAP: Positive Airway Pressure; NIV: Non-Invasive Ventilation.

At this version of the PM (version 1) the Protocol has been presented to Stakeholders and feedback is being acquired to update the Protocol in version 2, as per grant proposal. The specifications of the protocol for collection of the Pilot data will be documented in the D3.2 (Project Manual – CoBraD RWD, Updated MES-CoBraD Protocols and Quality and Quantity (Q&Q) Evaluation – v2) and based on integrated results from tasks T3.3 (Quality and Quantity Validation) and T3.4 (Methodology elaboration). The following appendices contain details on all current data acquired across sites, as well as guidelines on performing test runs for the next three months of the project. The integration of Stakeholder feedback

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

and assessment of test run successes and failures will form the final version of the harmonised Protocol in version 2.

ANNEX I : Data Abstraction Categories

Section 1 : H&P-based RWD categories

1. CC and HPI
2. ROS
3. FS / ADLs / IADLs
    a. Driving
    b. Cooking
    c. Exercise and Diet
    d. Financial management
    e. Self-care
4. PMH and Development
    a. ICD-x / DSM-y codes
    b. Review of high-yield medical history
    c. Review of diagnoses based on current and past medications
5. SH
    a. Location of birth and growing up
    b. Living and Relationship status
    c. Religion
    d. Education
    e. School / employment history and status
    f. Financial
    g. Healthcare System and Utilisation
6. FH
    a. Ancestry
    b. Family Social Background
    c. Family PMH
    d. Review of high-yield medical history (e.g., CoBraD)
7. Medications
    a. Current
        i. Prescribed
        ii. Day-to-day use
    b. Historical
        i. Review of medication class based on CCC (i.e., CoBraD)
8. Physical examination
    a. General
        i. Anthropometrics
    b. HENT
        i. Nasal
        ii. Oral-Pharyngeal
    c. Cardiac
    d. Respiratory
    e. Neurological
        i. Mental Status and Demeanor
        ii. Cranial Nerves
        iii. Strength and Reflexes
        iv. Sensation and Perception
        v. Coordination and Balance
9. Neuropsychological Testing

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

a. Regulatory (Alertness, Attention, Processing speed)
b. Memory
c. Language
d. Visuospatial
e. Executive

10. Laboratory testing
    a. Biosamples
    b. Imaging
    c. Neurophysiology

Section 2 : STS (Severity, Temporal, Spatial) categories

1. Severity
    a. Presence
    b. Severity
2. Chronic and Episodic Duration
    a. Chronic duration
    b. Episode duration
        i. Usual
        ii. Longest
3. Progression
    a. Gradual
    b. Stepwise
    c. Stable
    d. Relapsing Remitting
4. Frequency and Rhythmicity of Events
    a. Frequency
    b. Periodicity
5. Location
    a. Body part (Ocular, Axial, Arms, Legs)
    b. Symmetry
6. Symptom perception
    a. Positive vs Negative

Section 3 : Complex chronic condition-specific categories (CoBraD-specific categories)

1. Cognition and Behaviour
    a. Memory
        i. Short-term memory
        ii. Long-term memory
        iii. Visual memory
        iv. Verbal memory
    b. Executive
        i. Processing speed
        ii. Attention and alertness
        iii. Working memory and Multitasking
        iv. Planning and organisation
        v. Judgment and Complex Problem Solving
        vi. Cognitive Inhibition
        vii. Error monitoring
    c. Language and Speech

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

i. Naming
ii. Semantic comprehension
iii. Praxis
iv. Grammar
v. Prosody
vi. Intensity

d. Visual/Sensory Perceptual
i. Navigation
ii. Object Perception
iii. Movement Perception
iv. Auditory processing

e. Mood
i. Anxiety
ii. Sadness

f. Behaviour and Social Cognition
i. Social inhibition
ii. Empathy
iii. Motivation (vs. Apathy)
iv. Mental fluidity (vs. Rigidity and Perseveration)
v. Dietary habits (vs. Hyperorality)

2. Motor
a. Extrapyramidal
i. Bradykinesia and Freezing
ii. Rigidity
iii. Tremor
iv. Postural instability
v. Dystonia

b. Pyramidal
i. Limb-kinetic praxis
ii. UMN
iii. LMN

c. Praxis
i. Planning apraxia
ii. Chorea
iii. Cerebellar ataxia
iv. Myoclonus
v. Tardive dyskinesia

d. Perceptual
i. Sensory ataxia

3. Autonomic-Sensory-Perceptual
a. Temperature
b. Cardiovascular
c. Gastrointestinal
d. Olfaction
e. Audition
f. Vision
g. Somesthesia
h. Taste
i. Proprioception

4. Sleep

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

a. Sleep Disordered Breathing
b. Insomnia
    i. Sleep Onset
    ii. Sleep Maintenance
    iii. Early awakening
c. Hypersomnia and Hypersomnolence
    i. Excessive Daytime Somnolence
    ii. Excessive sleep time
d. Sleep-wake Rhythms
    i. Infradian
    ii. Circadian
    iii. Ultradian
e. Sleep-related movements
    i. Limb movements prior to sleep
f. Parasomnia and Dissociated State
    i. NREM parasomnia
    ii. REM parasomnia
    iii. REM-wake dissociation
    iv. NREM-wake dissociation

5. Epilepsy and Seizure Disorders
    a. Premonitory aura symptoms
        i. Character / Type
    b. Seizure
        i. Phenotype
            1. Focal vs Generalised
            2. Motor vs. Perceptual vs Autonomic
            3. Level of Alertness
        ii. Triggers
    c. Post-ictal
        i. Injuries
        ii. Cognitive / Behavioural deficits
        iii. Praxis deficits
        iv. Perceptual deficits
        v. Autonomic and Urinary deficits

Section 4 : Data Quality Category

1. Patient-caregiver effort and cooperation
2. Procedure quality

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

ANNEX II : Acquisition instruments for collecting clinical symptoms, history, physical exam and neuropsychology data

RWD category	Sub categories	Variable acquisition instruments
Clinical symptoms and history	Essential demographic , contact data and informant data	Part of intake questionnaire
	Social information	Employment, Income, Housing, Education years, Highest diploma, Religious Affinity, Developmental history, Childhood trauma,
	Healthcare data	Healthcare coverage/access to care/usual source of care
	Chief Complaint- Seizure	ASM Liverpool, SSQ
	Chief Complaint-Sleep	Sleep Diary, PSQI, ESS, ISI, STOP-BANG, BNSQ, MCTQ, rMEC
	Chief Complaint-Cognitive & Behavioural	NPI-Q, NPI-12, CDR
	Mood	GDS, HADS, PHQ-9, BDI, BAI
	Review of Systems (Neurological)	Olfacation, Incontinence,
	Disability catagories and questionnaires	General Disability (eyesight, hearing, motor, recent acute) , CDR, FAQ, QOLIE 35, TabCat BHS
	Habits	Smoking, Alcohol, Coffee, Illicit drug, Exercise, Diet
	Family History	As in appendix 1
	Neurological exam	As in appendix 1

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Physical Examination and Neuropsychological testing	Global Cognition	MMSE, MoCA, Addenbrook,
	Memory	CVLT short, CVLT standard, CERAD, WMS, FCRST, Sleep mediated memory consolidation, TabCat Favorites
	Executive	Digit Span, Phonemic Fluency, Category Fluency (animals / Fruits&Vegetables), 5 dots, Stroop, TMT A, TMT B, Modified Trails, Tabcat Match
	Visuospatial	Benson copy ,RCFT (<65), VOSP (numbers), CERAD (figures), Poppelreuter, Clocks, CATS , TabCat Line orientation
	Language	Boston naming 15 items,

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

ANNEX III : Neurophysiology Variables

EEG characteristics:

Electrode position , e.g standard 10-20 or Modified Maudsley system.

Electrode/contact impedances <5KΩ

Sampling rate of ≥256Hz.

Recording frequencies 0.1 Hz to 70Hz, with the 50 Hz filter OFF.

1 ECG left to right shoulder Low Frequency Filter (LF) 0.3 and High Frequency Filter (HF)70, sensitivity 500-1000 uV/mm.

Sleep recordings:

Electro-Oculo-Gram (EOG) electrodes

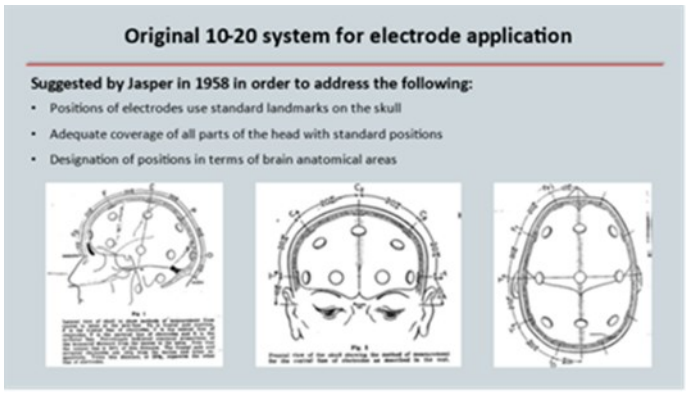
Right EOG-1cm above and lateral, Left EOG-1 cm below and lateral; LF 0.5Hz, HF 70Hz, Sen 5-10 Uv/mm

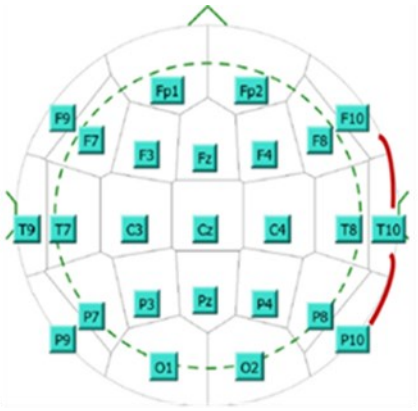
EMG electrodes (LH 10Hz, HF 100Hz) at face and muscular electrodes:

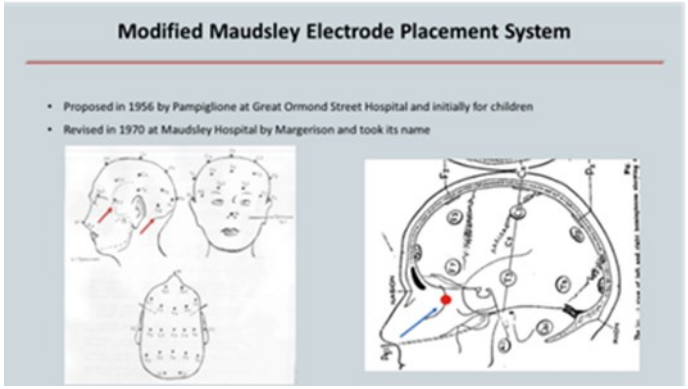
Chin, right and left deltoids, right and left Tibialis anterior RBD and oral (see above)

Respiratory belts for respiratory effort (chest and abdomen) and 1 Oral-nasal airflow LF 0.1Hz and HF 15Hz.

If possible, a sleep Score/ Apnoea /Hypopnoea Index (AHI)/ Arousal index could also be included in the data.







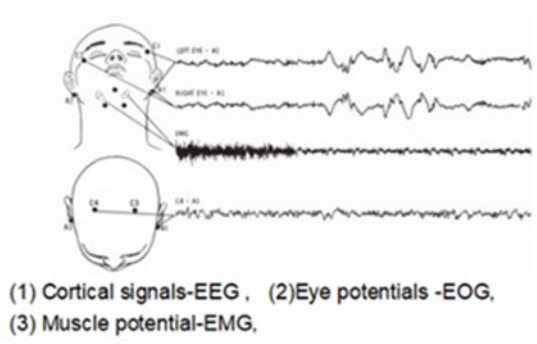


Figure 6 EEG Electrode Placement

ANNEX IV : Medical device RWD

VNS variables

1. Indication for VNS
2. VNS Model
3. Year of implantation
4. Dates of battery/Device replacement
5. Current parameters and previous parameters
    a. Intensity
    b. Frequency of stimulation
    c. Cycle On/Off
    d. Responsive stimulation parameters
6. Daily and summative machine-estimated outcome metrics:
    a. Number of response stimulations
7. Summary from the last VNS check

DBS variables

1. Indication for DBS
2. DBS Model
3. Year of implantation
4. Location of the leads
5. Pattern of leads' stimulation
6. Dates of battery/Device replacement
7. Current and previous parameters per period of use
    a. Intensity
    b. Frequency of stimulation
    c. Cycle
    d. Pulse width

PAP / NIV variables

1. Indication
2. Machine and Interface Model
3. Period of use parameters
    a. Pressure
    b. Humidity
    c. Temperature Parameters
4. Daily and summative machine-estimated outcome metrics:
    a. Use (hours and times)
    b. Apnea Hypopnea Index
    c. Leak

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

ANNEX V : Sampling Procedures flow chart

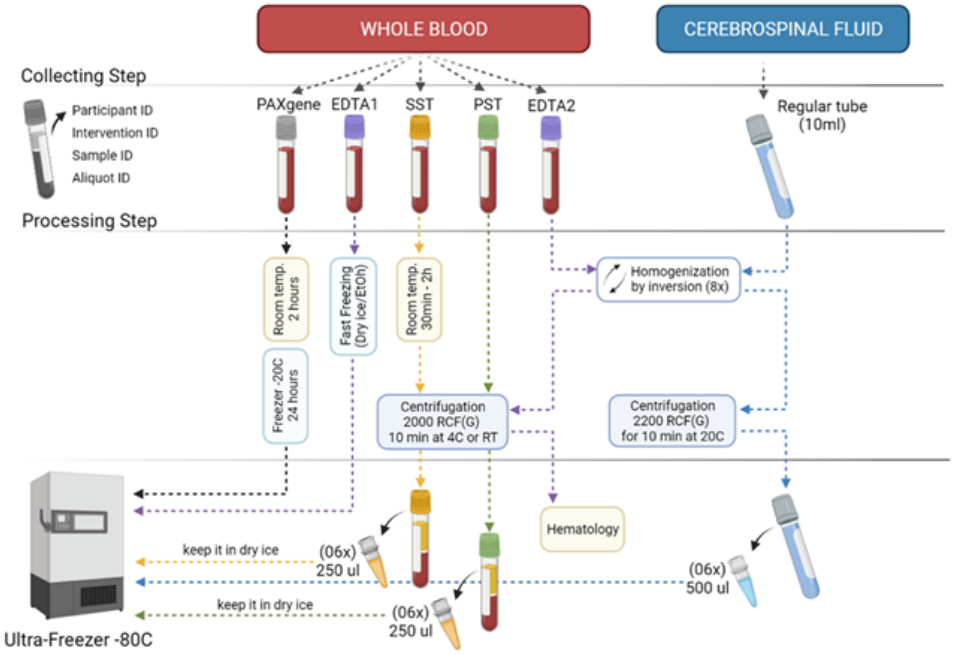


Figure 7 Sampling Procedure Flow chart. The gold standard for plasma samples is to be directly centrifuge after drawing. However, if that is not possible, a maximum of 90 minutes can be tolerated. However, this waiting time needs to be the same for all centers. So, we need to discuss more in the next meetings

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

ANNEX VI : Proteomic variables that can be measured from biological samples

Neurology Panel

Uniprot	Protein name	Gene name
Q9H2W6	39S ribosomal protein L46, mitochondrial	MRPL46
Q03393	6-pyruvoyl tetrahydrobiopterin synthase	PTS
P07108	Acyl-CoA-binding protein	DBI
P07741	Adenine phosphoribosyltransferase	APRT
O60242	Adhesion G protein-coupled receptor B3	ADGRB3
Q10588	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 2	BST1
Q9UJY5	ADP-ribosylation factor-binding protein GGA1	GGA1
Q9UKK9	ADP-sugar pyrophosphatase	NUDT5
Q9BYC5	Alpha-	FUT8
P30533	Alpha-2-macroglobulin receptor-associated protein	LRPAP1
P06733	Alpha-enolase	ENO1
P02771	Alpha-fetoprotein	AFP
Q9NZD4	Alpha-hemoglobin-stabilizing protein	AHSP
P40222	Alpha-taxilin	TXLNA
P05067	Amyloid-beta precursor protein	APP
Q9UJ72	Annexin A10	ANXA10
P12429	Annexin A3	ANXA3
P08758	Annexin A5	ANXA5
O95994	Anterior gradient protein 2 homolog	AGR2
Q07812	Apoptosis regulator BAX	BAX
P15289	Arylsulfatase A	ARSA
P07306	Asialoglycoprotein receptor 1	ASGR1
P14868	Aspartate--tRNA ligase, cytoplasmic	DARS1
Q9UBB4	Ataxin-10	ATXN10
P48047	ATP synthase subunit O, mitochondrial	ATP5PO
Q16740	ATP-dependent Clp protease proteolytic subunit, mitochondrial	CLPP
Q4LE39	AT-rich interactive domain-containing protein 4B	ARID4B
O95817	BAG family molecular chaperone regulator 3	BAG3
P50895	Basal cell adhesion molecule	BCAM
O43505	Beta-1,4-glucuronyltransferase 1	B4GAT1
P02749	Beta-2-glycoprotein 1	APOH
P16278	Beta-galactosidase	GLB1
Q86Z14	Beta-klotho	KLB
P01138	Beta-nerve growth factor	NGF
Q10589	Bone marrow stromal antigen 2	BST2
P12644	Bone morphogenetic protein 4	BMP4
Q96GW7	Brevican core protein	BCAN
Q9UBW5	Bridging integrator 2	BIN2
Q2TAL6	Brorin	VWC2
P55291	Cadherin-15	CDH15
P22223	Cadherin-3	CDH3
Q96JP9	Cadherin-related family member 1	CDHR1
P63098	Calcineurin subunit B type 1	PPP3R1
P01258	Calcitonin [Cleaved into: Calcitonin; Katacalcin	CALCA
Q9Y2V2	Calcium-regulated heat-stable protein 1	CARHSP1
P22676	Calretinin	CALB2
O94985	Calsyntenin-1	CLSTN1
P00918	Carbonic anhydrase 2	CA2
P23280	Carbonic anhydrase 6	CA6
P48052	Carboxypeptidase A2	CPA2
P14384	Carboxypeptidase M	CPM

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P29466	Caspase-1	CASP1
Q92851	Caspase-10	CASP10
P25774	Cathepsin S	CTSS
P11717	Cation-independent mannose-6-phosphate receptor	IGF2R
Q99731	C-C motif chemokine 19	CCL19
P13500	C-C motif chemokine 2	CCL2
O76076	CCN family member 5	CCN5
Q6YHK3	CD109 antigen	CD109
Q8N6Q3	CD177 antigen	CD177
P08962	CD63 antigen	CD63
P14209	CD99 antigen	CD99
Q8TCZ2	CD99 antigen-like protein 2	CD99L2
P41208	Centrin-2	CETN2
Q9Y5P4	Ceramide transfer protein	CERT
Q9HD42	Charged multivesicular body protein 1a	CHMP1A
P17538	Chymotrypsinogen B	CTRB1
Q9HAW4	Claspin	CLSPN
Q08708	CMRF35-like molecule 6	CD300C
Q6UXG3	CMRF35-like molecule 9	CD300LG
O14579	Coatomer subunit epsilon	COPE
Q6P1N0	Coiled-coil and C2 domain-containing protein 1A	CC2D1A
P04118	Colipase	CLPS
P45452	Collagenase 3	MMP13
Q9P232	Contactin-3	CNTN3
Q8IWV2	Contactin-4	CNTN4
O94779	Contactin-5	CNTN5
O14618	Copper chaperone for superoxide dismutase	CCS
Q9P126	C-type lectin domain family 1 member B	CLEC1B
Q8IUN9	C-type lectin domain family 10 member A	CLEC10A
Q9Y240	C-type lectin domain family 11 member A	CLEC11A
Q86T13	C-type lectin domain family 14 member A	CLEC14A
Q9H5V8	CUB domain-containing protein 1	CDCP1
O14625	C-X-C motif chemokine 11	CXCL11
O43927	C-X-C motif chemokine 13	CXCL13
P15336	Cyclic AMP-dependent transcription factor ATF-2	ATF2
P28325	Cystatin-D	CST5
P52943	Cysteine-rich protein 2	CRIP2
O95727	Cytotoxic and regulatory T-cell molecule	CRTAM
P78560	Death domain-containing protein CRADD	CRADD
Q14739	Delta-14-SR	LBR
Q02487	Desmocollin-2	DSC2
Q14126	Desmoglein-2	DSG2
O94907	Dickkopf-related protein 1	DKK1
Q9UBT3	Dickkopf-related protein 4	DKK4
P49789	Dinucleosidetriphosphatase	FHIT
P16444	Dipeptidase 1	DPEP1
Q9P0K1	Disintegrin and metalloproteinase domain-containing protein 22	ADAM22
P78325	Disintegrin and metalloproteinase domain-containing protein 8	ADAM8
Q13426	DNA repair protein XRCC4	XRCC4
Q8NBI3	Draxin	DRAXIN
P51452	Dual specificity protein phosphatase 3	DUSP3
O00399	Dynactin subunit 6	DCTN6
Q6XZF7	Dynamin-binding protein	DNMBP
Q9UKV5	E3 ubiquitin-protein ligase AMFR	AMFR
P14625	Endoplasmin	HSP90B1
P42892	Endothelin-converting enzyme 1	ECE1
Q5JZY3	Ephrin type-A receptor 10	EPHA10
O15197	Ephrin type-B receptor 6	EPHB6

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P20827	Ephrin-A1	EFNA1
P52798	Ephrin-A4	EFNA4
Q08345	Epithelial discoidin domain-containing receptor 1	DDR1
P23588	Eukaryotic translation initiation factor 4B	EIF4B
P15311	Ezrin	EZR
Q01469	Fatty acid-binding protein 5	FABP5
Q96RD9	Fc receptor-like protein 5	FCRL5
P30043	Flavin reductase	BLVRB
P14207	Folate receptor beta	FOLR2
O95466	Formin-like protein 1	FMNL1
P78423	Fractalkine	CX3CL1
Q9NQ88	Fructose-2,6-bisphosphatase TIGAR	TIGAR
O15117	FYN-binding protein 1	FYB1
P21217_Q11128	Galactoside 3_5	FUT3_FUT5
O00214	Galectin-8	LGALS8
P09104	Gamma-enolase	ENO2
O76070	Gamma-synuclein	SNCG
Q9NS71	Gastrokine-1	GKN1
O60609	GDNF family receptor alpha-3	GFRA3
P39905	Glial cell line-derived neurotrophic factor	GDNF
P19440	Glutathione hydrolase 1 proenzyme	GGT1
P36269	Glutathione hydrolase 5 proenzyme	GGT5
P09211	Glutathione S-transferase P	GSTP1
P09466	Glycodelin	PAEP
P01215	Glycoprotein hormones alpha chain	CGA
Q9H1C3	Glycosyltransferase 8 domain-containing protein 2	GLT8D2
P78333	Glypican-5 [Cleaved into: Secreted glypican-5]	GPC5
Q92917	G-patch domain and KOW motifs-containing protein	GPKOW
P15509	Granulocyte-macrophage colony-stimulating factor receptor subunit alpha	CSF2RA
P22749	Granulysin	GNLY
O15496	Group 10 secretory phospholipase A2	PLA2G10
Q02643	Growth hormone-releasing hormone receptor	GHRHR
Q96PP9	Guanylate-binding protein 4	GBP4
Q02747	Guanylin	GUCA2A
P28906	Hematopoietic progenitor cell antigen CD34	CD34
P30519	Heme oxygenase 2	HMOX2
Q8TDQ0	Hepatitis A virus cellular receptor 2	HAVCR2
P50135	Histamine N-methyltransferase	HNMT
P12081	Histidine--tRNA ligase, cytoplasmic	HARS1
Q53H47	Histone-lysine N-methyltransferase SETMAR	SETMAR
P04233	HLA class II histocompatibility antigen gamma chain	CD74
Q6UXK2	Immunoglobulin superfamily containing leucine-rich repeat protein 2	ISLR2
Q8NBJ7	Inactive C-alpha-formylglycine-generating enzyme 2	SUMF2
Q6UXH9	Inactive serine protease PAMR1	PAMR1
Q13308	Inactive tyrosine-protein kinase 7	PTK7
P55103	Inhibin beta C chain	INHBC
Q9UK53	Inhibitor of growth protein 1	ING1
P29218	Inositol monophosphatase 1	IMPA1
P22692	Insulin-like growth factor-binding protein 4	IGFBP4
P08648	Integrin alpha-5	ITGA5
P11215	Integrin alpha-M	ITGAM
Q9H0C8	Integrin-linked kinase-associated serine/threonine phosphatase 2C	ILKAP
Q8WWN9	Interactor protein for cytohesin exchange factors 1	IPCEF1
P38484	Interferon gamma receptor 2	IFNGR2
Q8IU54	Interferon lambda-1	IFNL1
Q9NPH3	Interleukin-1 receptor accessory protein	IL1RAP
P14778	Interleukin-1 receptor type 1	IL1R1
Q96F46	Interleukin-17 receptor A	IL17RA

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

O95256	Interleukin-18 receptor accessory protein	IL18RAP
Q6ZMJ4	Interleukin-34	IL34
P05231	Interleukin-6	IL6
P16871	Interleukin-7 receptor subunit alpha	IL7R
P10145	Interleukin-8	CXCL8
Q9BXS1	Isopentenyl-diphosphate delta-isomerase 2	IDI2
P26440	Isovaleryl-CoA dehydrogenase, mitochondrial	IVD
Q9Y624	Junctional adhesion molecule A	F11R
P57087	Junctional adhesion molecule B	JAM2
P23276	Kell blood group glycoprotein	KEL
P02533	Keratin, type I cytoskeletal 14	KRT14
P13647	Keratin, type II cytoskeletal 5	KRT5
Q6UWL6	Kin of IRRE-like protein 2	KIRREL2
O43278	Kunitz-type protease inhibitor 1	SPINT1
Q6PIL6	Kv channel-interacting protein 4	KCNIP4
Q08431	Lactadherin	MFGE8
P22079	Lactoperoxidase	LPO
Q9BS40	Latexin	LXN
Q6UX15	Layilin	LAYN
O43155	Leucine-rich repeat transmembrane protein FLRT2	FLRT2
P15018	Leukemia inhibitory factor	LIF
P30740	Leukocyte elastase inhibitor	SERPINB1
Q8N149	Leukocyte immunoglobulin-like receptor subfamily A member 2	LILRA2
Q6ISS4	Leukocyte-associated immunoglobulin-like receptor 2	LAIR2
Q6P1M0	Long-chain fatty acid transport protein 4	SLC27A4
P06734	Low affinity immunoglobulin epsilon Fc receptor	FCER2
Q14696	LRP chaperone MESD	MESD
P05455	Lupus La protein	SSB
Q8N2G4	Ly6/PLAUR domain-containing protein 1	LYPD1
Q9Y6Y9	Lymphocyte antigen 96	LY96
Q14108	Lysosome membrane protein 2	SCARB2
P13473	Lysosome-associated membrane glycoprotein 2	LAMP2
P14174	Macrophage migration inhibitory factor	MIF
P21757	Macrophage scavenger receptor types I and II	MSR1
Q8NFP4	MAM domain-containing glycosylphosphatidylinositol anchor protein 1	MDGA1
P48740	Mannan-binding lectin serine protease 1	MASP1
O15232	Matrilin-3	MATN3
P14780	Matrix metalloproteinase-9	MMP9
Q99727	Metalloproteinase inhibitor 4	TIMP4
P53582	Methionine aminopeptidase 1	METAP1
P10636	Microtubule-associated protein tau	MAPT
P20774	Mimecan	OGN
Q8WV92	MIT domain-containing protein 1	MITD1
Q9Y4K4	Mitogen-activated protein kinase kinase kinase kinase 5	MAP4K5
Q9Y6D9	Mitotic spindle assembly checkpoint protein MAD1	MAD1L1
P53985	Monocarboxylate transporter 1	SLC16A1
Q9H3R2	Mucin-13	MUC13
P05164	Myeloperoxidase	MPO
Q99972	Myocilin	MYOC
Q02083	N-acylethanolamine-hydrolyzing acid amidase	NAAA
Q9NXA8	NAD-dependent protein deacylase sirtuin-5, mitochondrial	SIRT5
O00308	NEDD4-like E3 ubiquitin-protein ligase WWP2	WWP2
O15394	Neural cell adhesion molecule 2	NCAM2
P58417	Neurexophilin-1	NXPH1
O14594	Neurocan core protein	NCAN
P07196	Neurofilament light polypeptide	NEFL
Q15818	Neuronal pentraxin-1	NPTX1
O60462	Neuropilin-2	NRP2

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q9NR71	Neutral ceramidase	ASAH2
P22894	Neutrophil collagenase	MMP8
Q9HAN9	Nicotinamide/nicotinic acid mononucleotide adenylyltransferase 1	NMNAT1
Q14112	Nidogen-2	NID2
P29475	Nitric oxide synthase, brain	NOS1
P29474	Nitric oxide synthase, endothelial	NOS3
Q15155	Nodal modulator 1	NOMO1
Q9UNZ2	NSFL1 cofactor p47	NSFL1C
Q16288	NT-3 growth factor receptor	NTRK3
P06748	Nucleophosmin	NPM1
A1E959	Odontogenic ameloblast-associated protein	ODAM
Q9NPH6	Odorant-binding protein 2b	OBP2B
P01178	Oxytocin-neurophysin 1	OXT
Q9UKJ1	Paired immunoglobulin-like type 2 receptor alpha	PILRA
Q99497	Parkinson disease protein 7	PARK7
P49023	Paxillin	PXN
Q02790	Peptidyl-prolyl cis-trans isomerase FKBP4	FKBP4
Q13451	Peptidyl-prolyl cis-trans isomerase FKBP5	FKBP5
Q9Y680	Peptidyl-prolyl cis-trans isomerase FKBP7	FKBP7
O60240	Perilipin-1	PLIN1
Q06830	Peroxiredoxin-1	PRDX1
P30039	Phenazine biosynthesis-like domain-containing protein	PBLD
P30086	Phosphatidylethanolamine-binding protein 1	PEBP1
P60484	Phosphatidylinositol 3,4,5-trisphosphate 3-phosphatase	PTEN
Q8TCT1	Phosphoethanolamine/phosphocholine phosphatase	PHOSPHO1
Q96FE7	Phosphoinositide-3-kinase-interacting protein 1	PIK3IP1
O14523	Phospholipid transfer protein C2CD2L	C2CD2L
Q15126	Phosphomevalonate kinase	PMVK
Q96CD2	Phosphopantothenoylcysteine decarboxylase	PPCDC
Q9NRG1	Phosphoribosyltransferase domain-containing protein 1	PRTFDC1
P16284	Platelet endothelial cell adhesion molecule	PECAM1
Q9HCN6	Platelet glycoprotein VI	GP6
Q13093	Platelet-activating factor acetylhydrolase	PLA2G7
P15151	Poliovirus receptor	PVR
P01833	Polymeric immunoglobulin receptor	PIGR
Q9UHV9	Prefoldin subunit 2	PFDN2
P11464	Pregnancy-specific beta-1-glycoprotein 1	PSG1
P51531	Probable global transcription activator SNF2L2	SMARCA2
O14944	Proepiregulin [Cleaved into: Epiregulin	EREG
Q9NZQ7	Programmed cell death 1 ligand 1	CD274
O14737	Programmed cell death protein 5	PDCD5
P28799	Progranulin	GRN
P01236	Prolactin	PRL
Q96B36	Proline-rich AKT1 substrate 1	AKT1S1
Q06323	Proteasome activator complex subunit 1	PSME1
Q9UL46	Proteasome activator complex subunit 2	PSME2
Q96IU4	Protein ABHD14B	ABHD14B
Q8WUW1	Protein BRICK1	BRK1
Q9NSK7	Protein C19orf12	C19orf12
P53539	Protein fosB	FOSB
Q6P4E1	Protein GOLM2	GOLM2
P61244	Protein max	MAX
Q92597	Protein NDRG1	NDRG1
Q96FQ6	Protein S100-A16	S100A16
Q9UM07	Protein-arginine deiminase type-4	PADI4
P20936	Ras GTPase-activating protein 1	RASA1
Q9NRW1	Ras-related protein Rab-6B	RAB6B
O00559	Receptor-binding cancer antigen expressed on SiSo cells	EBAG9

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q92932	Receptor-type tyrosine-protein phosphatase N2	PTPRN2
Q96B86	Repulsive guidance molecule A	RGMA
P00352	Retinal dehydrogenase 1	ALDH1A1
Q6NW40	RGM domain family member B	RGMB
P08134	Rho-related GTP-binding protein RhoC	RHOC
Q9H477	Ribokinase	RBKS
Q15633	RISC-loading complex subunit TARBP2	TARBP2
Q9HCK4	Roundabout homolog 2	ROBO2
Q2MKA7	R-spondin-1	RSPO1
Q9H446	RWD domain-containing protein 1	RWDD1
Q6ZMJ2	Scavenger receptor class A member 5	SCARA5
Q8WTV0	Scavenger receptor class B member 1	SCARB1
Q96GP6	Scavenger receptor class F member 2	SCARF2
Q86VW0	SEC14 domain and spectrin repeat-containing protein 1	SESTD1
Q8N474	Secreted frizzled-related protein 1	SFRP1
Q92765	Secreted frizzled-related protein 3	FRZB
P05060	Secretogranin-1	CHGB
Q92854	Semaphorin-4D	SEMA4D
P00995	Serine protease inhibitor Kazal-type 1	SPINK1
Q9NQ38	Serine protease inhibitor Kazal-type 5	SPINK5
Q9Y6E0	Serine/threonine-protein kinase 24	STK24
O96013	Serine/threonine-protein kinase PAK 4	PAK4
P37023	Serine/threonine-protein kinase receptor R3	ACVRL1
Q9BRF8	Serine/threonine-protein phosphatase CPPED1	CPPED1
P35237	Serpin B6	SERPINB6
P50453	Serpin B9	SERPINB9
Q6ZMC9	Sialic acid-binding Ig-like lectin 15	SIGLEC15
O15389	Sialic acid-binding Ig-like lectin 5	SIGLEC5
Q04900	Sialomucin core protein 24	CD164
O94813	Slit homolog 2 protein	SLIT2
P17405	Sphingomyelin phosphodiesterase	SMPD1
Q86WV1	Src kinase-associated phosphoprotein 1	SKAP1
O95630	STAM-binding protein	STAMBP
P52823	Stanniocalcin-1	STC1
O76061	Stanniocalcin-2	STC2
P31948	Stress-induced-phosphoprotein 1	STIP1
P08254	Stromelysin-1	MMP3
P50225	Sulfotransferase 1A1	SULT1A1
P04179	Superoxide dismutase [Mn], mitochondrial	SOD2
Q9UGT4	Sushi domain-containing protein 2	SUSD2
Q9HA65	TBC1 domain family member 17	TBC1D17
P56279	T-cell leukemia/lymphoma protein 1A	TCL1A
P01732	T-cell surface glycoprotein CD8 alpha chain	CD8A
Q92752	Tenascin-R	TNR
P22105	Tenascin-X	TNXB
P13385	Teratocarcinoma-derived growth factor 1	TDGF1
Q08629	Testican-1	SPOCK1
Q8NBS9	Thioredoxin domain-containing protein 5	TXNDC5
Q16881	Thioredoxin reductase 1, cytoplasmic	TXNRD1
Q16762	Thiosulfate sulfurtransferase	TST
P35442	Thrombospondin-2	THBS2
P04216	Thy-1 membrane glycoprotein	THY1
P63313	Thymosin beta-10	TMSB10
Q9H3S3	Transmembrane protease serine 5	TMPRSS5
P04155	Trefoil factor 1	TFF1
Q5T2D2	Trem-like transcript 2 protein	TREML2
P23381	Tryptophan--tRNA ligase, cytoplasmic	WARS
Q9BW30	Tubulin polymerization-promoting protein family member 3	TPPP3

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q99426	Tubulin-folding cofactor B	TBCB
Q15814	Tubulin-specific chaperone C	TBCC
Q9Y2W6	Tudor and KH domain-containing protein	TDRKH
P01375	Tumor necrosis factor	TNF
O43557	Tumor necrosis factor ligand superfamily member 14	TNFSF14
O00220	Tumor necrosis factor receptor superfamily member 10A	TNFRSF10A
O14763	Tumor necrosis factor receptor superfamily member 10B	TNFRSF10B
Q969Z4	Tumor necrosis factor receptor superfamily member 19L	RELT
P19438	Tumor necrosis factor receptor superfamily member 1A	TNFRSF1A
P20333	Tumor necrosis factor receptor superfamily member 1B	TNFRSF1B
O75509	Tumor necrosis factor receptor superfamily member 21	TNFRSF21
O95407	Tumor necrosis factor receptor superfamily member 6B	TNFRSF6B
P28908	Tumor necrosis factor receptor superfamily member 8	TNFRSF8
Q07011	Tumor necrosis factor receptor superfamily member 9	TNFRSF9
P09769	Tyrosine-protein kinase Fgr	FGR
P18031	Tyrosine-protein phosphatase non-receptor type 1	PTPN1
Q9BZM5	UL16-binding protein 2	ULBP2
P00749	Urokinase-type plasminogen activator	PLAU
Q9NP79	Vacuolar protein sorting-associated protein VTA1 homolog	VTA1
P13611	Versican core protein	VCAN
Q9Y279	V-set and immunoglobulin domain-containing protein 4	VSIG4
Q6UX27	V-set and transmembrane domain-containing protein 1	VSTM1
Q16864	V-type proton ATPase subunit F	ATP6V1F
Q96NZ8	WAP, Kazal, immunoglobulin, Kunitz and NTR domain-containing protein 1	WFIKKN1
Q9UPY6	Wiskott-Aldrich syndrome protein family member 3	WASF3
Q15043	Zinc transporter ZIP14	SLC39A14

Cardiometabolic Panel

Uniprot	Protein name	Gene name
Q76LX8	A disintegrin and metalloproteinase with thrombospondin motifs 13	ADAMTS13
Q8TE57	A disintegrin and metalloproteinase with thrombospondin motifs 16	ADAMTS16
P62736	Actin, aortic smooth muscle	ACTA2
Q9NZK5	Adenosine deaminase 2	ADA2
P23526	Adenosylhomocysteinase	AHCY
P00568	Adenylate kinase isoenzyme 1	AK1
P48960	Adhesion G protein-coupled receptor E5	ADGRE5
Q8IZP9	Adhesion G-protein coupled receptor G2	ADGRG2
P16112	Aggrecan core protein	ACAN
P17516	Aldo-keto reductase family 1 member C4	AKR1C4
P08319	All-trans-retinol dehydrogenase [NAD	ADH4
P19961	Alpha-amylase 2B	AMY2B
Q03154	Aminoacylase-1	ACY1
P15144	Aminopeptidase N	ANPEP
P51693	Amyloid-like protein 1	APLP1
P03950	Angiogenin	ANG
O95841	Angiopoietin-related protein 1	ANGPTL1
Q9Y5C1	Angiopoietin-related protein 3	ANGPTL3
Q9BYF1	Angiotensin-converting enzyme 2	ACE2
P09525	Annexin A4	ANXA4
O95445	Apolipoprotein M	APOM
Q9UBU3	Appetite-regulating hormone	GHRL
P20711	Aromatic-L-amino-acid decarboxylase	DDC
P20160	Azurocidin	AZU1
P98160	Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2
Q16620	BDNF/NT-3 growth factors receptor	NTRK2
Q96KN2	Beta-Ala-His dipeptidase	CNDP1
P15907	Beta-galactoside alpha-2,6-sialyltransferase 1	ST6GAL1
P08236	Beta-glucuronidase	GUSB

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P08118	Beta-microseminoprotein	MSMB
P34913	Bifunctional epoxide hydrolase 2 [Includes: Cytosolic epoxide hydrolase 2	EPHX2
Q13867	Bleomycin hydrolase	BLMH
P22004	Bone morphogenetic protein 6	BMP6
Q8TDL5	BPI fold-containing family B member 1	BPIFB1
Q9BWV1	Brother of CDO	BOC
P12830	Cadherin-1	CDH1
Q12864	Cadherin-17	CDH17
P19022	Cadherin-2	CDH2
P33151	Cadherin-5	CDH5
P55285	Cadherin-6	CDH6
Q9HBB8	Cadherin-related family member 5	CDHR5
P10644	cAMP-dependent protein kinase type I-alpha regulatory subunit	PRKAR1A
P00915	Carbonic anhydrase 1	CA1
Q8N1Q1	Carbonic anhydrase 13	CA13
P07451	Carbonic anhydrase 3	CA3
P22748	Carbonic anhydrase 4	CA4
P35218	Carbonic anhydrase 5A, mitochondrial	CA5A
P15085	Carboxypeptidase A1	CPA1
P15086	Carboxypeptidase B	CPB1
P31997	Carcinoembryonic antigen-related cell adhesion molecule 8	CEACAM8
Q16619	Cardiotrophin-1	CTF1
Q9NQ79	Cartilage acidic protein 1	CRTAC1
P49747	Cartilage oligomeric matrix protein	COMP
P42574	Caspase-3	CASP3
P21964	Catechol O-methyltransferase	COMT
P07858	Cathepsin B	CTSB
P07339	Cathepsin D	CTSD
P07711	Cathepsin L1	CTSL
Q9UBR2	Cathepsin Z	CTSZ
Q16627	C-C motif chemokine 14	CCL14
Q16663	C-C motif chemokine 15	CCL15
O15467	C-C motif chemokine 16	CCL16
P55774	C-C motif chemokine 18	CCL18
Q9Y4X3	C-C motif chemokine 27	CCL27
P13501	C-C motif chemokine 5	CCL5
P17676	CCAAT/enhancer-binding protein beta	CEBPB
P48745	CCN family member 3	CCN3
Q13740	CD166 antigen	ALCAM
Q9NNX6	CD209 antigen	CD209
Q9Y5K6	CD2-associated protein	CD2AP
P13987	CD59 glycoprotein	CD59
Q99674	Cell growth regulator with EF hand domain protein 1	CGREF1
P43121	Cell surface glycoprotein MUC18	MCAM
O95684	Centrosomal protein 43	CEP43
P36222	Chitinase-3-like protein 1	CHI3L1
Q13231	Chitotriosidase-1	CHIT1
Q6WN34	Chordin-like protein 2	CHRDL2
P09093	Chymotrypsin-like elastase family member 3A	CELA3A
P09496	Clathrin light chain A	CLTA
Q15846	Clusterin-like protein 1	CLUL1
P00740	Coagulation factor IX	F9
P08709	Coagulation factor VII	F7
P27352	Cobalamin binding intrinsic factor	CBLIF
Q76M96	Coiled-coil domain-containing protein 80	CCDC80
P02462	Collagen alpha-1	COL4A1
P39060	Collagen alpha-1	COL18A1
P02452	Collagen alpha-1	COL1A1

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P12111	Collagen alpha-3	COL6A3
Q9BXJ1	Complement C1q tumor necrosis factor-related protein 1	C1QTNF1
P06681	Complement C2	C2
Q9NPY3	Complement component C1q receptor	CD93
P08174	Complement decay-accelerating factor	CD55
P20023	Complement receptor type 2	CR2
O43186	Cone-rod homeobox protein	CRX
Q6PJW8	Consortin	CNST
Q12860	Contactin-1	CNTN1
P31146	Coronin-1A	CORO1A
P34998	Corticotropin-releasing factor receptor 1	CRHR1
Q8NC01	C-type lectin domain family 1 member A	CLEC1A
Q9NY25	C-type lectin domain family 5 member A	CLEC5A
Q9H2A7	C-X-C motif chemokine 16	CXCL16
P42830	C-X-C motif chemokine 5	CXCL5
P04080	Cystatin-B	CSTB
P01034	Cystatin-C	CST3
Q15828	Cystatin-M	CST6
Q9H773	dCTP pyrophosphatase 1	DCTPP1
P07585	Decorin	DCN
Q07507	Dermatopontin	DPT
Q9NR28	Diablo homolog, mitochondrial	DIABLO
Q9UBP4	Dickkopf-related protein 3	DKK3
P09417	Dihydropteridine reductase	QDPR
Q9UHL4	Dipeptidyl peptidase 2	DPP7
P27487	Dipeptidyl peptidase 4	DPP4
Q13444	Disintegrin and metalloproteinase domain-containing protein 15	ADAM15
Q8NHS0	DnaJ homolog subfamily B member 8	DNAJB8
O60496	Docking protein 2	DOK2
Q9NRD8	Dual oxidase 2	DUOX2
Q07108	Early activation antigen CD69	CD69
Q13508	Ecto-ADP-ribosyltransferase 3	ART3
O75356	Ectonucleoside triphosphate diphosphohydrolase 5	ENTPD5
O75354	Ectonucleoside triphosphate diphosphohydrolase 6	ENTPD6
Q13822	Ectonucleotide pyrophosphatase/phosphodiesterase family member 2	ENPP2
Q12805	EGF-containing fibulin-like extracellular matrix protein 1	EFEMP1
O43854	EGF-like repeat and discoidin I-like domain-containing protein 3	EDIL3
P19957	Elafin	PI3
P17813	Endoglin	ENG
Q53H82	Endoribonuclease LACTB2	LACTB2
Q96AP7	Endothelial cell-selective adhesion molecule	ESAM
P12724	Eosinophil cationic protein	RNASE3
P54760	Ephrin type-B receptor 4	EPHB4
P00533	Epidermal growth factor receptor	EGFR
P16581	E-selectin	SELE
Q13541	Eukaryotic translation initiation factor 4E-binding protein 1	EIF4EBP1
Q13158	FAS-associated death domain protein	FADD
P15090	Fatty acid-binding protein, adipocyte	FABP4
P12104	Fatty acid-binding protein, intestinal	FABP2
Q96LA6	Fc receptor-like protein 1	FCRL1
Q9UGM5	Fetuin-B	FETUB
Q15485	Ficolin-2	FCN2
P09467	Fructose-1,6-bisphosphatase 1	FBP1
P34947	G protein-coupled receptor kinase 5	GRK5
P09382	Galectin-1	LGALS1
Q05315	Galectin-10	CLC
P17931	Galectin-3	LGALS3
Q92820	Gamma-glutamyl hydrolase	GGH

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P51161	Gastrotropin	FABP6
Q16769	Glutaminyl-peptide cyclotransferase	QPCT
P35754	Glutaredoxin-1	GLRX
P08263	Glutathione S-transferase A1	GSTA1
P13807	Glycogen [starch] synthase, muscle	GYS1
P55808	Glycoprotein Xg	XG
P20718	Granzyme H	GZMH
075791	GRB2-related adapter protein 2	GRAP2
Q14393	Growth arrest-specific protein 6	GAS6
Q99988	Growth/differentiation factor 15	GDF15
Q9UK05	Growth/differentiation factor 2	GDF2
O14793	Growth/differentiation factor 8	MSTN
P04792	Heat shock protein beta-1	HSPB1
P09601	Heme oxygenase 1	HMOX1
Q9NRV9	Heme-binding protein 1	HEBP1
P08581	Hepatocyte growth factor receptor	MET
P61978	Heterogeneous nuclear ribonucleoprotein K	HNRNPK
P52789	Hexokinase-2	HK2
Q12794	Hyaluronidase-1	HYAL1
Q9Y4L1	Hypoxia up-regulated protein 1	HYOU1
Q969P0	Immunoglobulin superfamily member 8	IGSF8
Q01973	Inactive tyrosine-protein kinase transmembrane receptor ROR1	ROR1
Q12912	Inositol 1,4,5-triphosphate receptor associated 2	IRAG2
P08833	Insulin-like growth factor-binding protein 1	IGFBP1
P18065	Insulin-like growth factor-binding protein 2	IGFBP2
P17936	Insulin-like growth factor-binding protein 3	IGFBP3
P24592	Insulin-like growth factor-binding protein 6	IGFBP6
Q16270	Insulin-like growth factor-binding protein 7	IGFBP7
Q8WX77	Insulin-like growth factor-binding protein-like 1	IGFBPL1
P05556	Integrin beta-1	ITGB1
Q9UKP3	Integrin beta-1-binding protein 2	ITGB1BP2
P05107	Integrin beta-2	ITGB2
Q06033	Inter-alpha-trypsin inhibitor heavy chain H3	ITIH3
P05362	Intercellular adhesion molecule 1	ICAM1
P13598	Intercellular adhesion molecule 2	ICAM2
P32942	Intercellular adhesion molecule 3	ICAM3
Q9UMF0	Intercellular adhesion molecule 5	ICAM5
Q01638	Interleukin-1 receptor-like 1	IL1RL1
O95998	Interleukin-18-binding protein	IL18BP
Q9UHD0	Interleukin-19	IL19
P01589	Interleukin-2 receptor subunit alpha	IL2RA
P05231	Interleukin-6	IL6
P08887	Interleukin-6 receptor subunit alpha	IL6R
P40189	Interleukin-6 receptor subunit beta	IL6ST
P10145	Interleukin-8	CXCL8
P21583	Kit ligand	KITLG
Q16773	Kynurenine–oxoglutarate transaminase 1	KYAT1
Q04760	Lactoylglutathione lyase	GLO1
P46379	Large proline-rich protein BAG6	BAG6
Q14767	Latent-transforming growth factor beta-binding protein 2	LTBP2
P41159	Leptin	LEP
P48357	Leptin receptor	LEPR
A6NI73	Leukocyte immunoglobulin-like receptor subfamily A member 5	LILRA5
Q8NHL6	Leukocyte immunoglobulin-like receptor subfamily B member 1	LILRB1
Q8N423	Leukocyte immunoglobulin-like receptor subfamily B member 2	LILRB2
O75023	Leukocyte immunoglobulin-like receptor subfamily B member 5	LILRB5
P18428	Lipopolysaccharide-binding protein	LBP
P06858	Lipoprotein lipase	LPL

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P05451	Lithostathine-1-alpha	REG1A
P48304	Lithostathine-1-beta	REG1B
P23141	Liver carboxylesterase 1	CES1
P12318	Low affinity immunoglobulin gamma Fc region receptor II-a	FCGR2A
O75015	Low affinity immunoglobulin gamma Fc region receptor III-B	FCGR3B
P01130	Low-density lipoprotein receptor	LDLR
Q86VZ4	Low-density lipoprotein receptor-related protein 11	LRP11
P42785	Lysosomal Pro-X carboxypeptidase	PRCP
Q9UEW3	Macrophage receptor MARCO	MARCO
P10721	Mast/stem cell growth factor receptor Kit	KIT
P09237	Matrilysin	MMP7
P15529	Membrane cofactor protein	CD46
Q16853	Membrane primary amine oxidase	AOC3
Q16820	Meprin A subunit beta	MEP1B
P01033	Metalloproteinase inhibitor 1	TIMP1
P55082	Microfibril-associated glycoprotein 3	MFAP3
Q13361	Microfibrillar-associated protein 5	MFAP5
P08571	Monocyte differentiation antigen CD14	CD14
Q99549	M-phase phosphoprotein 8	MPHOSPH8
Q8NI22	Multiple coagulation factor deficiency protein 2	MCFD2
Q9H1U4	Multiple epidermal growth factor-like domains protein 9	MEGF9
P24158	Myeloblastin	PRTN3
P41218	Myeloid cell nuclear differentiation antigen	MNDA
P02144	Myoglobin	MB
P58546	Myotrophin	MTPN
O95544	NAD kinase	NADK
P16860	Natriuretic peptides B	NPPB
Q92692	Nectin-2	NECTIN2
P13591	Neural cell adhesion molecule 1	NCAM1
O00533	Neural cell adhesion molecule L1-like protein	CHL1
Q9NQX5	Neural proliferation differentiation and control protein 1	NPDC1
P46531	Neurogenic locus notch homolog protein 1	NOTCH1
Q9UM47	Neurogenic locus notch homolog protein 3	NOTCH3
Q92823	Neuronal cell adhesion molecule	NRCAM
O95502	Neuronal pentraxin receptor	NPTXR
O14786	Neuropilin-1	NRP1
P59665	Neutrophil defensin 1	DEFA1_DEFA1B
P80188	Neutrophil gelatinase-associated lipocalin	LCN2
Q6GTS8	N-fatty-acyl-amino acid synthase/hydrolase PM20D1	PM20D1
P14543	Nidogen-1	NID1
NT-proBNP	NT-proBNP	NTproBNP
P31483	Nucleolysin TIA-1 isoform p40	TIA1
Q99650	Oncostatin-M-specific receptor subunit beta	OSMR
P10451	Osteopontin	SPP1
P78380	Oxidized low-density lipoprotein receptor 1	OLR1
Q9UKJ0	Paired immunoglobulin-like type 2 receptor beta	PILRB
P04746	Pancreatic alpha-amylase	AMY2A
P55259	Pancreatic secretory granule membrane major glycoprotein GP2	GP2
O75594	Peptidoglycan recognition protein 1	PGLYRP1
P19021	Peptidyl-glycine alpha-amidating monooxygenase	PAM
P23284	Peptidyl-prolyl cis-trans isomerase B	PPIB
O60664	Perilipin-3	PLIN3
Q15067	Peroxisomal acyl-coenzyme A oxidase 1	ACOX1
P04054	Phospholipase A2	PLA2G1B
P14555	Phospholipase A2, membrane associated	PLA2G2A
P55058	Phospholipid transfer protein	PLTP
Q9NWQ8	Phosphoprotein associated with glycosphingolipid-enriched microdomains 1	PAG1
P05121	Plasminogen activator inhibitor 1	SERPINE1

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q5VY43	Platelet endothelial aggregation receptor 1	PEAR1
P07359	Platelet glycoprotein Ib alpha chain	GP1BA
P16234	Platelet-derived growth factor receptor alpha	PDGFRA
P09619	Platelet-derived growth factor receptor beta	PDGFRB
P04085	Platelet-derived growth factor subunit A	PDGFA
P21246	Pleiotrophin	PTN
O15031	Plexin-B2	PLXNB2
Q9ULL4	Plexin-B3	PLXNB3
P09668	Pro-cathepsin H [Cleaved into: Cathepsin H mini chain; Cathepsin H	CTSH
Q15113	Procollagen C-endopeptidase enhancer 1	PCOLCE
O75340	Programmed cell death protein 6	PDCD6
Q12884	Prolyl endopeptidase FAP	FAP
Q8NBP7	Proprotein convertase subtilisin/kexin type 9	PCSK9
O15354	Prosaposin receptor GPR37	GPR37
P41222	Prostaglandin-H2 D-isomerase	PTGDS
Q9Y2B0	Protein canopy homolog 2	CNPY2
P80370	Protein delta homolog 1	DLK1
Q92520	Protein FAM3C	FAM3C
P41236	Protein phosphatase inhibitor 2	PPP1R2
P31949	Protein S100-A11	S100A11
P25815	Protein S100-P	S100P
P21980	Protein-glutamine gamma-glutamyltransferase 2	TGM2
O14917	Protocadherin-17	PCDH17
P16109	P-selectin	SELP
P35247	Pulmonary surfactant-associated protein D	SFTPD
O94903	Pyridoxal phosphate homeostasis protein	PLPBP
P10586	Receptor-type tyrosine-protein phosphatase F	PTPRF
Q13332	Receptor-type tyrosine-protein phosphatase S	PTPRS
Q06141	Regenerating islet-derived protein 3-alpha	REG3A
P00797	Renin	REN
Q9HD89	Resistin	RETN
Q9UKL0	REST corepressor 1	RCOR1
Q99969	Retinoic acid receptor responder protein 2	RARRES2
O00584	Ribonuclease T2	RNASET2
Q14162	Scavenger receptor class F member 1	SCARF1
Q8WTU2	Scavenger receptor cysteine-rich domain-containing group B protein	SSC4D
Q86VB7	Scavenger receptor cysteine-rich type 1 protein M130	CD163
Q9BQB4	Sclerostin	SOST
Q13275	Semaphorin-3F	SEMA3F
O75326	Semaphorin-7A	SEMA7A
Q9BQR3	Serine protease 27	PRSS27
Q13043	Serine/threonine-protein kinase 4	STK4
O96017	Serine/threonine-protein kinase Chk2	CHEK2
Q15831	Serine/threonine-protein kinase STK11	STK11
P21549	Serine--pyruvate aminotransferase	AGXT
Q86U17	Serpin A11	SERPINA11
Q8IW75	Serpin A12	SERPINA12
P36952	Serpin B5	SERPINB5
Q15165	Serum paraoxonase/arylesterase 2	PON2
Q9Y286	Sialic acid-binding Ig-like lectin 7	SIGLEC7
Q9H5Y7	SLIT and NTRK-like protein 6	SLITRK6
Q8WVQ1	Soluble calcium-activated nucleotidase 1	CANT1
A1L4H1	Soluble scavenger receptor cysteine-rich domain-containing protein SSC5D	SSC5D
P01241	Somatotropin	GH1
Q99523	Sortilin	SORT1
Q9Y5X1	Sorting nexin-9	SNX9
Q14515	SPARC-like protein 1	SPARCL1
Q9BUD6	Spondin-2	SPON2

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P00441	Superoxide dismutase [Cu-Zn]	SOD1
Q6UWL2	Sushi domain-containing protein 1	SUSD1
O00161	Synaptosomal-associated protein 23	SNAP23
P18827	Syndecan-1	SDC1
P31431	Syndecan-4	SDC4
P13686	Tartrate-resistant acid phosphatase type 5	ACP5
Q96H15	T-cell immunoglobulin and mucin domain-containing protein 4	TIMD4
O95988	T-cell leukemia/lymphoma protein 1B	TCL1B
P24821	Tenascin	TNC
O60635	Tetraspanin-1	TSPAN1
P52888	Thimet oligopeptidase	THOP1
P07204	Thrombomodulin	THBD
P40225	Thrombopoietin	THPO
P35443	Thrombospondin-4	THBS4
Q969D9	Thymic stromal lymphopoietin	TSLP
P19971	Thymidine phosphorylase	TYMP
P01222	Thyrotropin subunit beta	TSHB
P04066	Tissue alpha-L-fucosidase	FUCA1
P10646	Tissue factor pathway inhibitor	TFPI
P00750	Tissue-type plasminogen activator	PLAT
P20062	Transcobalamin-2	TCN2
P02786	Transferrin receptor protein 1	TFRC
Q03167	Transforming growth factor beta receptor type 3	TGFBR3
Q15582	Transforming growth factor-beta-induced protein ig-h3	TGFBI
Q14956	Transmembrane glycoprotein NMB	GPNMB
Q07654	Trefoil factor 3	TFF3
P19429	Troponin I, cardiac muscle	TNNI3
P07478	Trypsin-2	PRSS2
Q9GZM7	Tubulointerstitial nephritis antigen-like	TINAGL1
P01375	Tumor necrosis factor	TNF
Q9Y275	Tumor necrosis factor ligand superfamily member 13B	TNFSF13B
O14798	Tumor necrosis factor receptor superfamily member 10C	TNFRSF10C
P25445	Tumor necrosis factor receptor superfamily member 6	FAS
Q96A56	Tumor protein p53-inducible nuclear protein 1	TP53INP1
P35590	Tyrosine-protein kinase receptor Tie-1	TIE1
Q06418	Tyrosine-protein kinase receptor TYRO3	TYRO3
P30530	Tyrosine-protein kinase receptor UFO	AXL
P78324	Tyrosine-protein phosphatase non-receptor type substrate 1	SIRPA
P40818	Ubiquitin carboxyl-terminal hydrolase 8	USP8
P07911	Uromodulin	UMOD
P19320	Vascular cell adhesion protein 1	VCAM1
Q6EMK4	Vasorin	VASN
O95183	Vesicle-associated membrane protein 5	VAMP5
P08670	Vimentin	VIM
P04070	Vitamin K-dependent protein C	PROC
P04275	von Willebrand factor	VWF
Q96N03	V-set and transmembrane domain-containing protein 2-like protein	VSTM2L
Q9H7M9	V-type immunoglobulin domain-containing suppressor of T-cell activation	VSIR
Q92558	Wiskott-Aldrich syndrome protein family member 1	WASF1
Q13105	Zinc finger and BTB domain-containing protein 17	ZBTB17

Inflammation Panel

Uniprot	Protein name	Gene name
Q16698	2,4-dienoyl-CoA reductase, mitochondrial	DECR1
O43598	2'-deoxynucleoside 5'-phosphate N-hydrolase 1	DNPH1
Q9Y478	5'-AMP-activated protein kinase subunit beta-1	PRKAB1
Q92484	Acid sphingomyelinase-like phosphodiesterase 3a	SMPDL3A
P00813	Adenosine deaminase	ADA

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q9UHX3	Adhesion G protein-coupled receptor E2	ADGRE2
Q15109	Advanced glycosylation end product-specific receptor	AGER
O00253	Agouti-related protein	AGRP
O00468	Agrin	AGRN
P30838	Aldehyde dehydrogenase, dimeric NADP-preferring	ALDH3A1
Q7Z6M3	Allergin-1	MILR1
O43707	Alpha-actinin-4	ACTN4
Q9NP70	Ameloblastin	AMBN
P19801	Amiloride-sensitive amine oxidase [copper-containing]	AOC1
Q15389	Angiopoietin-1	ANGPT1
Q9UKU9	Angiopoietin-related protein 2	ANGPTL2
Q9BY76	Angiopoietin-related protein 4	ANGPTL4
P50995	Annexin A11	ANXA11
Q5T4W7	Artemin	ARTN
P27540	Aryl hydrocarbon receptor nuclear translocator	ARNT
Q9UII2	ATPase inhibitor, mitochondrial	ATP5IF1
O15169	Axin-1	AXIN1
P35613	Basigin	BSG
P40259	B-cell antigen receptor complex-associated protein beta chain	CD79B
P20273	B-cell receptor CD22	CD22
Q8NDB2	B-cell scaffold protein with ankyrin repeats	BANK1
O43521-2	Bcl-2-like protein 11, Isoform BimL	BCL2L11
P15291	Beta-1,4-galactosyltransferase 1	B4GALT1
P55957	BH3-interacting domain death agonist	BID
Q06520	Bile salt sulfotransferase	SULT2A1
P11274	Breakpoint cluster region protein	BCR
Q7KYR7	Butyrophilin subfamily 2 member A1	BTN2A1
P78410	Butyrophilin subfamily 3 member A2	BTN3A2
Q9H4D0	Calsyntenin-2	CLSTN2
Q9UDT6	CAP-Gly domain-containing linker protein 2	CLIP2
Q3KPI0	Carcinoembryonic antigen-related cell adhesion molecule 21	CEACAM21
P42575	Caspase-2	CASP2
P43234	Cathepsin O	CTSO
Q99616	C-C motif chemokine 13	CCL13
Q92583	C-C motif chemokine 17	CCL17
P78556	C-C motif chemokine 20	CCL20
O00585	C-C motif chemokine 21	CCL21
O00626	C-C motif chemokine 22	CCL22
P55773	C-C motif chemokine 23	CCL23
O00175	C-C motif chemokine 24	CCL24
O15444	C-C motif chemokine 25	CCL25
Q9Y258	C-C motif chemokine 26	CCL26
Q9NRJ3	C-C motif chemokine 28	CCL28
P10147	C-C motif chemokine 3	CCL3
P13236	C-C motif chemokine 4	CCL4
P80098	C-C motif chemokine 7	CCL7
P29279	CCN family member 2	CCN2
O95971	CD160 antigen	CD160
Q5ZPR3	CD276 antigen	CD276
P29965	CD40 ligand	CD40LG
P09326	CD48 antigen	CD48
P32970	CD70 antigen	CD70
Q01151	CD83 antigen	CD83
Q4KMG0	Cell adhesion molecule-related/down-regulated by oncogenes	CDON
Q8TD46	Cell surface glycoprotein CD200 receptor 1	CD200R1
Q9UPV0	Centrosomal protein of 164 kDa	CEP164
Q9BU40	Chordin-like protein 1	CHRDL1
Q99895	Chymotrypsin-C	CTRC

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P20849	Collagen alpha-1	COL9A1
Q5KU26	Collectin-12	COLEC12
P02745	Complement C1q subcomponent subunit A	C1QA
Q9UHC6	Contactin-associated protein-like 2	CNTNAP2
Q15517	Corneodesmosin	CDSN
P28845	Corticosteroid 11-beta-dehydrogenase isozyme 1	HSD11B1
P24387	Corticotropin-releasing factor-binding protein	CRHBP
P78310	Coxsackievirus and adenovirus receptor	CXADR
P12532	Creatine kinase U-type, mitochondrial	CKMT1A_CKMT1B
P46109	Crk-like protein	CRKL
Q9UMR7	C-type lectin domain family 4 member A	CLEC4A
Q8WTT0	C-type lectin domain family 4 member C	CLEC4C
Q8WXI8	C-type lectin domain family 4 member D	CLEC4D
Q6UXB4	C-type lectin domain family 4 member G	CLEC4G
Q9BXN2	C-type lectin domain family 7 member A	CLEC7A
P23582	C-type natriuretic peptide [Cleaved into: CNP-22; CNP-29; CNP-53]	NPPC
P02778	C-X-C motif chemokine 10	CXCL10
O95715	C-X-C motif chemokine 14	CXCL14
Q6UXB2	C-X-C motif chemokine 17	CXCL17
P19876	C-X-C motif chemokine 3	CXCL3
P80162	C-X-C motif chemokine 6	CXCL6
Q07325	C-X-C motif chemokine 9	CXCL9
O76096	Cystatin-F	CST7
Q9NZV1	Cysteine-rich motor neuron 1 protein	CRIM1
O75462	Cytokine receptor-like factor 1	CRLF1
O43639	Cytoplasmic protein NCK2	NCK2
Q07065	Cytoskeleton-associated protein 4	CKAP4
P28838	Cytosol aminopeptidase	LAP3
Q9H0P0	Cytosolic 5'-nucleotidase 3A	NT5C3A
P47712	Cytosolic phospholipase A2	PLA2G4A
Q8NFT8	Delta and Notch-like epidermal growth factor-related receptor	DNER
Q13574	Diacylglycerol kinase zeta	DGKZ
P53634	Dipeptidyl peptidase 1	CTSC
O75077	Disintegrin and metalloproteinase domain-containing protein 23	ADAM23
O00273	DNA fragmentation factor subunit alpha	DFFA
O60884	DnaJ homolog subfamily A member 2	DNAJA2
Q9UJU6	Drebrin-like protein	DBNL
Q9UN19	Dual adapter for phosphotyrosine and 3-phosphotyrosine and 3-phosphoinositide	DAPP1
P52564	Dual specificity mitogen-activated protein kinase kinase 6	MAP2K6
Q14118	Dystroglycan	DAG1
P19474	E3 ubiquitin-protein ligase TRIM21	TRIM21
Q9UJA9	Ectonucleotide pyrophosphatase/phosphodiesterase family member 5	ENPP5
Q6UWV6	Ectonucleotide pyrophosphatase/phosphodiesterase family member 7	ENPP7
Q9GZT9	Egl nine homolog 1	EGLN1
Q9NQ30	Endothelial cell-specific molecule 1	ESM1
P51671	Eotaxin	CCL11
P21709	Ephrin type-A receptor 1	EPHA1
P16422	Epithelial cell adhesion molecule	EPCAM
P01588	Erythropoietin	EPO
Q04637	Eukaryotic translation initiation factor 4 gamma 1	EIF4G1
P63241	Eukaryotic translation initiation factor 5A-1	EIF5A
Q0Z7S8	Fatty acid-binding protein 9	FABP9
P07148	Fatty acid-binding protein, liver	FABP1
Q96LA5	Fc receptor-like protein 2	FCRL2
Q96P31	Fc receptor-like protein 3	FCRL3
Q6DN72	Fc receptor-like protein 6	FCRL6
O95750	Fibroblast growth factor 19	FGF19
P09038	Fibroblast growth factor 2	FGF2

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P12034	Fibroblast growth factor 5	FGF5
P49771	Fms-related tyrosine kinase 3 ligand	FLT3LG
P19883	Follistatin	FST
O95633	Follistatin-related protein 3	FSTL3
Q12778	Forkhead box protein O1	FOXO1
Q96DB9	FXYD domain-containing ion transport regulator 5	FXYD5
P22466	Galanin peptides [Cleaved into: Galanin; Galanin message-associated peptide	GAL
P56470	Galectin-4	LGALS4
O00182	Galectin-9	LGALS9
Q9HC38	Glyoxalase domain-containing protein 4	GLOD4
P36959	GMP reductase 1	GMPR
Q9HD26	Golgi-associated PDZ and coiled-coil motif-containing protein	GOPC
P09919	Granulocyte colony-stimulating factor	CSF3
P12544	Granzyme A	GZMA
P10144	Granzyme B	GZMB
P09341	Growth-regulated alpha protein	CXCL1
P32456	Guanylate-binding protein 2	GBP2
P0DMV8	Heat shock 70 kDa protein 1A	HSPA1A
P14317	Hematopoietic lineage cell-specific protein	HCLS1
P14210	Hepatocyte growth factor	HGF
P37235	Hippocalcin-like protein 1	HPCAL1
P13747	HLA class I histocompatibility antigen, alpha chain E	HLA-E
P01903	HLA class II histocompatibility antigen, DR alpha chain	HLA-DRA
P22304	Iduronate 2-sulfatase	IDS
P24071	Immunoglobulin alpha Fc receptor	FCAR
P01591	Immunoglobulin J chain	JCHAIN
Q8N608	Inactive dipeptidyl peptidase 10	DPP10
O43736	Integral membrane protein 2A	ITM2A
Q9UKX5	Integrin alpha-11	ITGA11
P23229	Integrin alpha-6	ITGA6
P18564	Integrin beta-6	ITGB6
Q14773	Intercellular adhesion molecule 4	ICAM4
P01579	Interferon gamma	IFNG
P15260	Interferon gamma receptor 1	IFNGR1
Q8IU57	Interferon lambda receptor 1	IFNLR1
P01583	Interleukin-1 alpha	IL1A
P01584	Interleukin-1 beta	IL1B
P18510	Interleukin-1 receptor antagonist protein	IL1RN
P27930	Interleukin-1 receptor type 2	IL1R2
P51617	Interleukin-1 receptor-associated kinase 1	IRAK1
Q9NWZ3	Interleukin-1 receptor-associated kinase 4	IRAK4
Q9HB29	Interleukin-1 receptor-like 2	IL1RL2
P22301	Interleukin-10	IL10
Q13651	Interleukin-10 receptor subunit alpha	IL10RA
Q08334	Interleukin-10 receptor subunit beta	IL10RB
P20809	Interleukin-11	IL11
P42701	Interleukin-12 receptor subunit beta-1	IL12RB1
P29460	Interleukin-12 subunit beta	IL12B
P35225	Interleukin-13	IL13
P40933	Interleukin-15	IL15
Q13261	Interleukin-15 receptor subunit alpha	IL15RA
Q9NRM6	Interleukin-17 receptor B	IL17RB
Q16552	Interleukin-17A	IL17A
Q9P0M4	Interleukin-17C	IL17C
Q8TAD2	Interleukin-17D	IL17D
Q96PD4	Interleukin-17F	IL17F
Q14116	Interleukin-18	IL18
Q13478	Interleukin-18 receptor 1	IL18R1

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P60568	Interleukin-2	IL2
P14784	Interleukin-2 receptor subunit beta	IL2RB
Q9NYY1	Interleukin-20	IL20
Q9UHF4	Interleukin-20 receptor subunit alpha	IL20RA
Q8N6P7	Interleukin-22 receptor subunit alpha-1	IL22RA1
Q13007	Interleukin-24	IL24
P26951	Interleukin-3 receptor subunit alpha	IL3RA
P24001	Interleukin-32	IL32
O95760	Interleukin-33	IL33
P05112	Interleukin-4	IL4
P24394	Interleukin-4 receptor subunit alpha	IL4R
P05113	Interleukin-5	IL5
Q01344	Interleukin-5 receptor subunit alpha	IL5RA
P05231	Interleukin-6	IL6
P13232	Interleukin-7	IL7
P10145	Interleukin-8	CXCL8
P03956	Interstitial collagenase	MMP1
Q05084	Islet cell autoantigen 1	ICA1
B1AKI9	Isthmin-1	ISM1
P08727	Keratin, type I cytoskeletal 19	KRT19
Q12918	Killer cell lectin-like receptor subfamily B member 1	KLRB1
O43291	Kunitz-type protease inhibitor 2	SPINT2
Q16719	Kynureninase	KYNU
Q16363	Laminin subunit alpha-4	LAMA4
Q99538	Legumain	LGMN
Q6UXK5	Leucine-rich repeat neuronal protein 1	LRRN1
P42702	Leukemia inhibitory factor receptor	LIFR
Q8NHJ6	Leukocyte immunoglobulin-like receptor subfamily B member 4	LILRB4
Q6GTX8	Leukocyte-associated immunoglobulin-like receptor 1	LAIR1
O43561	Linker for activation of T-cells family member 1	LAT
Q14210	Lymphocyte antigen 6D	LY6D
O60449	Lymphocyte antigen 75	LY75
P19256	Lymphocyte function-associated antigen 3	CD58
P33241	Lymphocyte-specific protein 1	LSP1
P01374	Lymphotoxin-alpha	LTA
Q9UQV4	Lysosome-associated membrane glycoprotein 3	LAMP3
P09603	Macrophage colony-stimulating factor 1	CSF1
Q8WU39	Marginal zone B- and B1-cell-specific protein	MZB1
O00339	Matrilin-2	MATN2
Q9NQ76	Matrix extracellular phosphoglycoprotein	MEPE
O95866	Megakaryocyte and platelet inhibitory receptor G6b	MPIG6B
P55145	Mesencephalic astrocyte-derived neurotrophic factor	MANF
P35625	Metalloproteinase inhibitor 3	TIMP3
Q6UB28	Methionine aminopeptidase 1D, mitochondrial	METAP1D
P16455	Methylated-DNA--protein-cysteine methyltransferase	MGMT
Q03426	Mevalonate kinase	MVK
Q29980_Q29983	MHC class I polypeptide-related sequence B_A	MICB_MICA
Q9Y3D6	Mitochondrial fission 1 protein	FIS1
P45984	Mitogen-activated protein kinase 9	MAPK9
Q99685	Monoglyceride lipase	MGLL
Q96KG7	Multiple epidermal growth factor-like domains protein 10	MEGF10
O76036	Natural cytotoxicity triggering receptor 1	NCR1
Q9BZW8	Natural killer cell receptor 2B4	CD244
Q13241	Natural killer cells antigen CD94	KLRD1
Q9Y5A7	NEDD8 ultimate buster 1	NUB1
Q96SB3	Neurabin-2	PPP1R9B
O94856	Neurofascin	NFASC
P20783	Neurotrophin-3	NTF3

Q99748	Neurturin	NRTN
P19878	Neutrophil cytosol factor 2	NCF2
Q9Y6K9	NF-kappa-B essential modulator	IKBKG
O60934	Nibrin	NBN
Q969V3	Nicalin	NCLN
O95644	Nuclear factor of activated T-cells, cytoplasmic 1	NFATC1
Q12968	Nuclear factor of activated T-cells, cytoplasmic 3	NFATC3
Q9Y266	Nuclear migration protein nudC	NUDC
Q13232	Nucleoside diphosphate kinase 3	NME3
P13725	Oncostatin-M	OSM
Q8IYS5	Osteoclast-associated immunoglobulin-like receptor	OSCAR
Q99983	Osteomodulin	OMD
P41217	OX-2 membrane glycoprotein	CD200
P54317	Pancreatic lipase-related protein 2	PNLIPRP2
Q13219	Pappalysin-1	PAPPA
Q03431	Parathyroid hormone/parathyroid hormone-related peptide receptor	PTH1R
O75475	PC4 and SFRS1-interacting protein	PSIP1
Q9NR12	PDZ and LIM domain protein 7	PDLIM7
P26022	Pentraxin-related protein PTX3	PTX3
P68106	Peptidyl-prolyl cis-trans isomerase FKBP1B	FKBP1B
P30044	Peroxiredoxin-5, mitochondrial	PRDX5
O60542	Persephin	PSPN
Q6ZUJ8	Phosphoinositide 3-kinase adapter protein 1	PIK3AP1
Q9H008	Phospholysine phosphohistidine inorganic pyrophosphate phosphatase	LHPP
P49763	Placenta growth factor	PGF
P01127	Platelet-derived growth factor subunit B	PDGFB
Q9HCM2	Plexin-A4	PLXNA4
P09874	Poly [ADP-ribose] polymerase 1	PARP1
Q14435	Polypeptide N-acetylgalactosaminyltransferase 3	GALNT3
Q8TCS8	Polyribonucleotide nucleotidyltransferase 1, mitochondrial	PNPT1
P01133	Pro-epidermal growth factor	EGF
Q14005	Pro-interleukin-16 [Cleaved into: Interleukin-16	IL16
P58294	Prokineticin-1	PROK1
Q9HCU5	Prolactin regulatory element-binding protein	PREB
P51888	Prolargin	PRELP
P12872	Promotilin [Cleaved into: Motilin; Motilin-associated peptide	MLN
Q16651	Prostasin	PRSS8
Q9BT73	Proteasome assembly chaperone 3	PSMG3
Q9BXJ7	Protein amnionless [Cleaved into: Soluble protein amnionless]	AMN
Q6UXH1	Protein disulfide isomerase CRELD2	CRELD2
Q8N8S7	Protein enabled homolog	ENAH
O94992	Protein HEXIM1	HEXIM1
Q04759	Protein kinase C theta type	PRKCQ
Q99435	Protein kinase C-binding protein NELL2	NELL2
Q8WV07	Protein LTO1 homolog	LTO1
O43597	Protein sprouty homolog 2	SPRY2
O14904	Protein Wnt-9a	WNT9A
Q9Y2J8	Protein-arginine deiminase type-2	PADI2
P25116	Proteinase-activated receptor 1	F2R
Q08174	Protocadherin-1	PCDH1
P01135	Protransforming growth factor alpha [Cleaved into: Transforming growth factor alpha	TGFA
Q14242	P-selectin glycoprotein ligand 1	SELPLG
P30613	Pyruvate kinase PKLR	PKLR
Q5R372	Rab GTPase-activating protein 1-like	RABGAP1L
Q96AX2	Ras-related protein Rab-37	RAB37
P20340	Ras-related protein Rab-6A	RAB6A
P21860	Receptor tyrosine-protein kinase erbB-3	ERBB3
P28827	Receptor-type tyrosine-protein phosphatase mu	PTPRM

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q9BYZ8	Regenerating islet-derived protein 4	REG4
P57771	Regulator of G-protein signaling 8	RGS8
Q9NZN5	Rho guanine nucleotide exchange factor 12	ARHGEF12
Q9Y6N7	Roundabout homolog 1	ROBO1
Q12765	Secernin-1	SCRN1
O76038	Secretagogin	SCGN
Q96PL1	Secretoglobin family 3A member 2	SCGB3A2
Q8WXD2	Secretogranin-3	SCG3
P34896	Serine hydroxymethyltransferase, cytosolic	SHMT1
O60575	Serine protease inhibitor Kazal-type 4	SPINK4
P50452	Serpin B8	SERPINB8
Q15166	Serum paraoxonase/lactonase 3	PON3
O60880	SH2 domain-containing protein 1A	SH2D1A
Q96LC7	Sialic acid-binding Ig-like lectin 10	SIGLEC10
Q9BZZ2	Sialoadhesin	SIGLEC1
Q13291	Signaling lymphocytic activation molecule	SLAMF1
Q9Y3P8	Signaling threshold-regulating transmembrane adapter 1	SIT1
O00241	Signal-regulatory protein beta-1	SIRPB1
Q9UIB8	SLAM family member 5	CD84
Q9NQ25	SLAM family member 7	SLAMF7
Q9H3U7	SPARC-related modular calcium-binding protein 2	SMOC2
Q9HCB6	Spondin-1	SPON1
O75563	Src kinase-associated phosphoprotein 2	SKAP2
P78362	SRSF protein kinase 2	SRPK2
Q8IVG5	Sterile alpha motif domain-containing protein 9-like	SAMD9L
P48061	Stromal cell-derived factor 1	CXCL12
P09238	Stromelysin-2	MMP10
Q9UNK0	Syntaxin-8	STX8
Q92609	TBC1 domain family member 5	TBC1D5
P30203	T-cell differentiation antigen CD6	CD6
P01730	T-cell surface glycoprotein CD4	CD4
P30048	Thioredoxin-dependent peroxide reductase, mitochondrial	PRDX3
Q9HBG7	T-lymphocyte surface antigen Ly-9	LY9
Q12933	TNF receptor-associated factor 2	TRAF2
O15455	Toll-like receptor 3	TLR3
Q92844	TRAF family member-associated NF-kappa-B activator	TANK
P05412	Transcription factor AP-1	JUN
O14867	Transcription regulator protein BACH1	BACH1
P01137	Transforming growth factor beta-1 proprotein [Cleaved into: Latency-associated peptide	TGFB1
P13693	Translationally-controlled tumor protein	TPT1
Q03403	Trefoil factor 2	TFF2
Q9NZC2	Triggering receptor expressed on myeloid cells 2	TREM2
Q9C035	Tripartite motif-containing protein 5	TRIM5
O14773	Tripeptidyl-peptidase 1	TPP1
Q15661	Tryptase alpha/beta-1	TPSAB1
Q7L8A9	Tubulinyl-Tyr carboxypeptidase 1	VASH1
P01375	Tumor necrosis factor	TNF
O95379	Tumor necrosis factor alpha-induced protein 8	TNFAIP8
P50591	Tumor necrosis factor ligand superfamily member 10	TNFSF10
O14788	Tumor necrosis factor ligand superfamily member 11	TNFSF11
O43508	Tumor necrosis factor ligand superfamily member 12	TNFSF12
O75888	Tumor necrosis factor ligand superfamily member 13	TNFSF13
P48023	Tumor necrosis factor ligand superfamily member 6	FASLG
Q9Y6Q6	Tumor necrosis factor receptor superfamily member 11A	TNFRSF11A
O00300	Tumor necrosis factor receptor superfamily member 11B	TNFRSF11B
O14836	Tumor necrosis factor receptor superfamily member 13B	TNFRSF13B
Q96RJ3	Tumor necrosis factor receptor superfamily member 13C	TNFRSF13C
Q92956	Tumor necrosis factor receptor superfamily member 14	TNFRSF14

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P36941	Tumor necrosis factor receptor superfamily member 3	LTBR
P43489	Tumor necrosis factor receptor superfamily member 4	TNFRSF4
P25942	Tumor necrosis factor receptor superfamily member 5	CD40
Q9UNE0	Tumor necrosis factor receptor superfamily member EDAR	EDAR
Q12866	Tyrosine-protein kinase Mer	MERTK
P29350	Tyrosine-protein phosphatase non-receptor type 6	PTPN6
Q13459	Unconventional myosin-IXb	MYO9B
Q03405	Urokinase plasminogen activator surface receptor	PLAUR
P11684	Uteroglobin	SCGB1A1
P15692	Vascular endothelial growth factor A	VEGFA
O43915	Vascular endothelial growth factor D	VEGFD
Q8TEU8	WAP, Kazal, immunoglobulin, Kunitz and NTR domain-containing protein 2	WFIKKN2
P42768	Wiskott-Aldrich syndrome protein	WAS
Q7Z739	YTH domain-containing family protein 3	YTHDF3
Q6ZMH5	Zinc transporter ZIP5	SLC39A5

Oncology Panel

Uniprot	Protein name	Gene name
O00233	26S proteasome non-ATPase regulatory subunit 9	PSMD9
P09110	3-ketoacyl-CoA thiolase, peroxisomal	ACAA1
P21589	5'-nucleotidase	NT5E
O60825	6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 2	PFKFB2
Q8TE58	A disintegrin and metalloproteinase with thrombospondin motifs 15	ADAMTS15
Q9UP79	A disintegrin and metalloproteinase with thrombospondin motifs 8	ADAMTS8
Q9Y653	Adhesion G-protein coupled receptor G1	ADGRG1
P28907	ADP-ribosyl cyclase/cyclic ADP-ribose hydrolase 1	CD38
Q9BTE6	Alanyl-tRNA editing protein Aarsd1	AARSD1
P15121	Aldo-keto reductase family 1 member B1	AKR1B1
P05187	Alkaline phosphatase, placental type	ALPP
P55008	Allograft inflammatory factor 1	AIF1
P35475	Alpha-L-iduronidase	IDUA
P15514	Amphiregulin	AREG
Q86SJ2	Amphoterin-induced protein 2	AMIGO2
Q7Z5R6	Amyloid beta A4 precursor protein-binding family B member 1-interacting protein	APBB1IP
Q6FI81	Anamorsin	CIAPIN1
Q02763	Angiopoietin-1 receptor	TEK
O15123	Angiopoietin-2	ANGPT2
O43827	Angiopoietin-related protein 7	ANGPTL7
Q6NXT1	Ankyrin repeat domain-containing protein 54	ANKRD54
Q8TD06	Anterior gradient protein 3	AGR3
O95786	Antiviral innate immune response receptor RIG-I	DDX58
O95831	Apoptosis-inducing factor 1, mitochondrial	AIFM1
P05089	Arginase-1	ARG1
P15848	Arylsulfatase B	ARSB
Q13490	Baculoviral IAP repeat-containing protein 2	BIRC2
Q02742	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase	GCNT1
O15263	Beta-defensin 4A	DEFB4A_DEFB4B
Q9Y223	Bifunctional UDP-N-acetylglucosamine 2-epimerase/N-acetylmannosamine kinase	GNE
P21810	Biglycan	BGN
Q13145	BMP and activin membrane-bound inhibitor homolog	BAMBI
Q9UQB8	Brain-specific angiogenesis inhibitor 1-associated protein 2	BAIAP2
P20851	C4b-binding protein beta chain	C4BPB
Q9BYE9	Cadherin-related family member 2	CDHR2
P05937	Calbindin	CALB1
Q8N5S9	Calcium/calmodulin-dependent protein kinase kinase 1	CAMKK1
Q9P1Z2	Calcium-binding and coiled-coil domain-containing protein 1	CALCOCO1
O43570	Carbonic anhydrase 12	CA12
Q9ULX7	Carbonic anhydrase 14	CA14

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q16790	Carbonic anhydrase 9	CA9
O75493	Carbonic anhydrase-related protein 11	CA11
Q6UWW8	Carboxylesterase 3	CES3
P16870	Carboxypeptidase E	CPE
P13688	Carcinoembryonic antigen-related cell adhesion molecule 1	CEACAM1
P40198	Carcinoembryonic antigen-related cell adhesion molecule 3	CEACAM3
P06731	Carcinoembryonic antigen-related cell adhesion molecule 5	CEACAM5
Q14790	Caspase-8	CASP8
Q9UBX1	Cathepsin F	CTSF
O60911	Cathepsin L2	CTSV
P80075	C-C motif chemokine 8	CCL8
O00622	CCN family member 1	CCN1
O95388	CCN family member 4	CCN4
P26842	CD27 antigen	CD27
Q8IX05	CD302 antigen	CD302
P30260	Cell division cycle protein 27 homolog	CDC27
Q9NX58	Cell growth-regulating nucleolar protein	LYAR
Q99795	Cell surface A33 antigen	GPA33
P04637	Cellular tumor antigen p53	TP53
Q96NB1	Centrosomal protein 20	CEP20
Q6P2H3	Centrosomal protein of 85 kDa	CEP85
Q9NTU7	Cerebellin-4	CBLN4
Q49AH0	Cerebral dopamine neurotrophic factor	CDNF
Q7Z5A7	Chemokine-like protein TAFA-5	TAFA5
Q11201	CMP-N-acetylneuraminate-beta-galactosamide-alpha-2,3-sialyltransferase 1	ST3GAL1
Q8TDQ1	CMRF35-like molecule 1	CD300LF
Q496F6	CMRF35-like molecule 2	CD300E
O00748	Cocaine esterase	CES2
Q02246	Contactin-2	CNTN2
O00244	Copper transport protein ATOX1	ATOX1
Q9UBG3	Cornulin	CRNN
P06850	Corticoliberin	CRH
P0CG37	Cryptic protein	CFC1
Q9UJ71	C-type lectin domain family 4 member K	CD207
Q6EIG7	C-type lectin domain family 6 member A	CLEC6A
Q9H6B4	CXADR-like membrane protein	CLMP
P55273	Cyclin-dependent kinase 4 inhibitor D	CDKN2D
P38936	Cyclin-dependent kinase inhibitor 1	CDKN1A
Q8WYN0	Cysteine protease ATG4A	ATG4A
P16562	Cysteine-rich secretory protein 2	CRISP2
P10606	Cytochrome c oxidase subunit 5B, mitochondrial	COX5B
O00548	Delta-like protein 1	DLL1
P32926	Desmoglein-3	DSG3
Q86SJ6	Desmoglein-4	DSG4
P55039	Developmentally-regulated GTP-binding protein 2	DRG2
P50583	Diadenosine tetraphosphatase	NUDT2
Q9UK85	Dickkopf-like protein 1	DKKL1
O94760	Dimethylarginine dimethylaminohydrolase 1	DDAH1
Q9H4A9	Dipeptidase 2	DPEP2
P42658	Dipeptidyl aminopeptidase-like protein 6	DPP6
P98082	Disabled homolog 2	DAB2
Q96PD2	Discoidin, CUB and LCCL domain-containing protein 2	DCBLD2
P27695	DNA-(apurinic or apyrimidinic site) endonuclease	APEX1
P61218	DNA-directed RNA polymerases I, II, and III subunit RPABC2	POLR2F
P25685	DnaJ homolog subfamily B member 1	DNAJB1
Q14203	Dynactin subunit 1	DCTN1
Q13561	Dynactin subunit 2	DCTN2
Q9H4P4	E3 ubiquitin-protein ligase NRDP1	RNF41

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q7L5Y9	E3 ubiquitin-protein transferase MAEA	MAEA
Q9Y5L3	Ectonucleoside triphosphate diphosphohydrolase 2	ENTPD2
Q9BSW2	EF-hand calcium-binding domain-containing protein 4B	CRACR2A
Q14241	Elongin-A	ELOA
Q9BS26	Endoplasmic reticulum resident protein 44	ERP44
P98073	Enteropeptidase	TMPRSS15
P29317	Ephrin type-A receptor 2	EPHA2
Q9H6S3	Epidermal growth factor receptor kinase substrate 8-like protein 2	EPS8L2
Q9UHF1	Epidermal growth factor-like protein 7	EGFL7
Q96RT1	Erbin	ERBIN
P55789	FAD-linked sulfhydryl oxidase ALR	GFER
O60907	F-box-like/WD repeat-containing protein TBL1X	TBL1X
Q6BAA4	Fc receptor-like B	FCRLB
Q9NSA1	Fibroblast growth factor 21	FGF21
Q9GZV9	Fibroblast growth factor 23	FGF23
P21802	Fibroblast growth factor receptor 2	FGFR2
Q14512	Fibroblast growth factor-binding protein 1	FGFBP1
P39748	Flap endonuclease 1	FEN1
P15328	Folate receptor alpha	FOLR1
P41439	Folate receptor gamma	FOLR3
O43524	Forkhead box protein O3	FOXO3
Q16595	Frataxin, mitochondrial	FXN
Q01543	Friend leukemia integration 1 transcription factor	FLI1
P09958	Furin	FURIN
P47929	Galectin-7	LGALS7_LGALS7B
A4D1B5	Gamma-secretase-activating protein	GSAP
P56159	GDNF family receptor alpha-1	GFRA1
O00451	GDNF family receptor alpha-2	GFRA2
O60763	General vesicular transport factor p115	USO1
P14136	Glial fibrillary acidic protein	GFAP
Q3B7J2	Glucose-fructose oxidoreductase domain-containing protein 2	GFOD2
Q16772	Glutathione S-transferase A3	GSTA3
Q8WUX2	Glutathione-specific gamma-glutamylcyclotransferase 2	CHAC2
P35052	Glypican-1 [Cleaved into: Secreted glypican-1]	GPC1
P01242	Growth hormone variant	GH2
Q9HAV7	GrpE protein homolog 1, mitochondrial	GRPEL1
O14558	Heat shock protein beta-6	HSPB6
O60760	Hematopoietic prostaglandin D synthase	HPGDS
P09105	Hemoglobin subunit theta-1	HBQ1
Q9Y662	Heparan sulfate glucosamine 3-O-sulfotransferase 3B1	HS3ST3B1
O60243	Heparan-sulfate 6-O-sulfotransferase 1	HS6ST1
Q96D42	Hepatitis A virus cellular receptor 1	HAVCR1
O14964	Hepatocyte growth factor-regulated tyrosine kinase substrate	HGS
P51858	Hepatoma-derived growth factor	HDGF
Q16543	Hsp90 co-chaperone Cdc37	CDC37
Q9UJM8	Hydroxyacid oxidase 1	HAO1
Q16775	Hydroxyacylglutathione hydrolase, mitochondrial	HAGH
O75144	ICOS ligand	ICOSLG
O75054	Immunoglobulin superfamily member 3	IGSF3
P49441	Inositol polyphosphate 1-phosphatase	INPP1
P08069	Insulin-like growth factor 1 receptor	IGF1R
P06756	Integrin alpha-V	ITGAV
O14713	Integrin beta-1-binding protein 1	ITGB1BP1
P18084	Integrin beta-5	ITGB5
P26010	Integrin beta-7	ITGB7
O75569	Interferon-inducible double-stranded RNA-dependent protein kinase activator A	PRKRA
P29459_P29460	Interleukin-12	IL12A_IL12B
P78552	Interleukin-13 receptor subunit alpha-1	IL13RA1

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q14213_Q8NEV9	Interleukin-27	EBI3_IL27
P05231	Interleukin-6	IL6
P10145	Interleukin-8	CXCL8
P06870	Kallikrein-1	KLK1
O43240	Kallikrein-10	KLK10
Q9UBX7	Kallikrein-11	KLK11
Q9UKR0	Kallikrein-12	KLK12
Q9UKR3	Kallikrein-13	KLK13
Q9P0G3	Kallikrein-14	KLK14
Q9Y5K2	Kallikrein-4	KLK4
Q92876	Kallikrein-6	KLK6
O60259	Kallikrein-8	KLK8
Q96I82	Kazal-type serine protease inhibitor domain-containing protein 1	KAZALD1
P05783	Keratin, type I cytoskeletal 18	KRT18
Q96EK5	KIF-binding protein	KIFBP
P43628	Killer cell immunoglobulin-like receptor 2DL3	KIR2DL3
P43629	Killer cell immunoglobulin-like receptor 3DL1	KIR3DL1
Q9NS15	Latent-transforming growth factor beta-binding protein 3	LTBP3
O00292	Left-right determination factor 2	LEFTY2
Q8N386	Leucine-rich repeat-containing protein 25	LRRC25
Q96JA1	Leucine-rich repeats and immunoglobulin-like domains protein 1	LRIG1
P09960	Leukotriene A-4 hydrolase	LTA4H
Q9GZY6	Linker for activation of T-cells family member 2	LAT2
P31994	Low affinity immunoglobulin gamma Fc region receptor II-b	FCGR2B
P01229	Lutropin subunit beta	LHB
Q7Z4W1	L-xylulose reductase	DCXR
O95274	Ly6/PLAUR domain-containing protein 3	LYPD3
Q6UX82	Ly6/PLAUR domain-containing protein 8	LYPD8
P18627	Lymphocyte activation gene 3 protein	LAG3
P47992	Lymphotactin	XCL1
Q9NPH0	Lysophosphatidic acid phosphatase type 6	ACP6
Q7L5N7	Lysophosphatidylcholine acyltransferase 2	LPCAT2
P39900	Macrophage metalloelastase	MMP12
P40121	Macrophage-capping protein	CAPG
P34949	Mannose-6-phosphate isomerase	MPI
Q9H8J5	MANSC domain-containing protein 1	MANSC1
Q9BUE0	Mediator of RNA polymerase II transcription subunit 18	MED18
Q9Y5V3	Melanoma-associated antigen D1	MAGED1
Q16674	Melanoma-derived growth regulatory protein	MIA
Q13421	Mesothelin	MSLN
P50579	Methionine aminopeptidase 2	METAP2
P21741	Midkine	MDK
Q7Z434	Mitochondrial antiviral-signaling protein	MAVS
Q9UJ68	Mitochondrial peptide methionine sulfoxide reductase	MSRA
Q99683	Mitogen-activated protein kinase kinase kinase 5	MAP3K5
Q08AG7	Mitotic-spindle organising protein 1	MZT1
Q15797	Mothers against decapentaplegic homolog 1	SMAD1
Q99717	Mothers against decapentaplegic homolog 5	SMAD5
Q8WXI7	Mucin-16	MUC16
Q16653	Myelin-oligodendrocyte glycoprotein	MOG
P20138	Myeloid cell surface antigen CD33	CD33
Q86SF2	N-acetylgalactosaminyltransferase 7	GALNT7
Q8IXJ6	NAD-dependent protein deacetylase sirtuin-2	SIRT2
O75380	NADH dehydrogenase [ubiquinone] iron-sulfur protein 6, mitochondrial	NDUFS6
Q96NY8	Nectin-4	NECTIN4
P08473	Neprilysin	MME
P32004	Neural cell adhesion molecule L1	L1CAM
P41271	Neuroblastoma suppressor of tumorigenicity 1	NBL1

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

P62166	Neuronal calcium sensor 1	NCS1
Q9Y639	Neuroplastin	NPTN
P34130	Neurotrophin-4	NTF4
O00221	NF-kappa-B inhibitor epsilon	NFKBIE
P43490	Nicotinamide phosphoribosyltransferase	NAMPT
Q92982	Ninjurin-1	NINJ1
P22307	Non-specific lipid-transfer protein	SCP2
P80303	Nucleobindin-2	NUCB2
P23515	Oligodendrocyte-myelin glycoprotein	OMG
Q9NZT2	Opioid growth factor receptor	OGFR
Q9UBM4	Opticin	OPTC
P01298	Pancreatic prohormone	PPY
P20472	Parvalbumin alpha	PVALB
P30041	Peroxiredoxin-6	PRDX6
O15357	Phosphatidylinositol 3,4,5-trisphosphate 5-phosphatase 2	INPPL1
Q8NCC3	Phospholipase A2 group XV	PLA2G15
P41586	Pituitary adenylate cyclase-activating polypeptide type I receptor	ADCYAP1R1
Q9NRA1	Platelet-derived growth factor C	PDGFC
Q8IUK5	Plexin domain-containing protein 1	PLXDC1
O00592	Podocalyxin	PODXL
Q9NZ53	Podocalyxin-like protein 2	PODXL2
O60828	Polyglutamine-binding protein 1	PQBP1
Q86SR1	Polypeptide N-acetylgalactosaminyltransferase 10	GALNT10
Q10471	Polypeptide N-acetylgalactosaminyltransferase 2	GALNT2
P08397	Porphobilinogen deaminase	HMBS
P35318	Pro-adrenomedullin	ADM
Q96SM3	Probable carboxypeptidase X1	CPXM1
Q8N9I9	Probable E3 ubiquitin-protein ligase DTX3	DTX3
Q9H3G5	Probable serine carboxypeptidase CPVL	CPVL
P35070	Probetacellulin [Cleaved into: Betacellulin	BTC
P01275	Pro-glucagon	GCG
Q9BQ51	Programmed cell death 1 ligand 2	PDCD1LG2
Q15116	Programmed cell death protein 1	PDCD1
Q99075	Proheparin-binding EGF-like growth factor [Cleaved into: Heparin-binding EGF-like growth factor	HBEGF
Q6PGN9	Proline/serine-rich coiled-coil protein 1	PSRC1
Q07954	Prolow-density lipoprotein receptor-related protein 1	LRP1
P01303	Pro-neuropeptide Y [Cleaved into: Neuropeptide Y	NPY
P25786	Proteasome subunit alpha type-1	PSMA1
P02760	Protein AMBP [Cleaved into: Alpha-1-microglobulin	AMBP
Q8N129	Protein canopy homolog 4	CNPY4
O75629	Protein CREG1	CREG1
P07237	Protein disulfide-isomerase	P4HB
Q9C005	Protein dpy-30 homolog	DPY30
P58499	Protein FAM3B	FAM3B
Q92832	Protein kinase C-binding protein NELL1	NELL1
O14974	Protein phosphatase 1 regulatory subunit 12A	PPP1R12A
P35813	Protein phosphatase 1A	PPM1A
Q9Y570	Protein phosphatase methylesterase 1	PPME1
P80511	Protein S100-A12	S100A12
P26447	Protein S100-A4	S100A4
Q9Y2Z0	Protein SGT1 homolog	SUGT1
O75695	Protein XRP2	RP2
Q2VWP7	Protogenin	PRTG
P07949	Proto-oncogene tyrosine-protein kinase receptor Ret	RET
P12931	Proto-oncogene tyrosine-protein kinase Src	SRC
Q8IWL2	Pulmonary surfactant-associated protein A1	SFTPA1
Q8IWL1	Pulmonary surfactant-associated protein A2	SFTPA2

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

Q9P0J1	Pyruvate dehydrogenase phosphatase catalytic subunit 1	PDP1
Q7Z6M1	Rab9 effector protein with kelch motifs	RABEPK
Q96NA2	Rab-interacting lysosomal protein	RILP
Q9Y243	RAC-gamma serine/threonine-protein kinase	AKT3
P46060	Ran GTPase-activating protein 1	RANGAP1
P50749	Ras association domain-containing protein 2	RASSF2
Q13576	Ras GTPase-activating-like protein IQGAP2	IQGAP2
P04626	Receptor tyrosine-protein kinase erbB-2	ERBB2
Q15303	Receptor tyrosine-protein kinase erbB-4	ERBB4
P36888	Receptor-type tyrosine-protein kinase FLT3	FLT3
O75787	Renin receptor	ATP6AP2
Q9BSG5	Retbindin	RTBDN
Q9BZR6	Reticulon-4 receptor	RTN4R
P49788	Retinoic acid receptor responder protein 1	RARRES1
P50120	Retinol-binding protein 2	RBP2
P82980	Retinol-binding protein 5	RBP5
Q07960	Rho GTPase-activating protein 1	ARHGAP1
P42331	Rho GTPase-activating protein 25	ARHGAP25
P31350	Ribonucleoside-diphosphate reductase subunit M2	RRM2
Q7LG56	Ribonucleoside-diphosphate reductase subunit M2 B	RRM2B
P35637	RNA-binding protein FUS	FUS
Q9BXY4	R-spondin-3	RSPO3
Q9Y265	RuvB-like 1	RUVBL1
P13521	Secretogranin-2	SCG2
O14828	Secretory carrier-associated membrane protein 3	SCAMP3
Q9BYH1	Seizure 6-like protein	SEZ6L
Q6UXD5	Seizure 6-like protein 2	SEZ6L2
Q96I15	Selenocysteine lyase	SCLY
Q9C0C4	Semaphorin-4C	SEMA4C
Q9UHD8	Septin-9	SEPTIN9
O43464	Serine protease HTRA2, mitochondrial	HTRA2
Q6UWN8	Serine protease inhibitor Kazal-type 6	SPINK6
Q86WD7	Serpin A9	SERPINA9
Q9UQQ2	SH2B adapter protein 3	SH2B3
Q9HAT2	Sialate O-acetylesterase	SIAE
O43699	Sialic acid-binding Ig-like lectin 6	SIGLEC6
Q9Y336	Sialic acid-binding Ig-like lectin 9	SIGLEC9
P37108	Signal recognition particle 14 kDa protein	SRP14
P51692	Signal transducer and activator of transcription 5B	STAT5B
Q96DU3	SLAM family member 6	SLAMF6
Q9P0V8	SLAM family member 8	SLAMF8
Q9H156	SLIT and NTRK-like protein 2	SLITRK2
Q00796	Sorbitol dehydrogenase	SORD
P09486	SPARC	SPARC
Q9H4F8	SPARC-related modular calcium-binding protein 1	SMOC1
Q15427	Splicing factor 3B subunit 4	SF3B4
Q06787	Synaptic functional regulator FMR1	FMR1
Q99536	Synaptic vesicle membrane protein VAT-1 homolog	VAT1
O95721	Synaptosomal-associated protein 29	SNAP29
O14662	Syntaxin-16	STX16
Q12846	Syntaxin-4	STX4
O43752	Syntaxin-6	STX6
O00186	Syntaxin-binding protein 3	STXBP3
Q9NUY8	TBC1 domain family member 23	TBC1D23
P29017	T-cell surface glycoprotein CD1c	CD1C
P06127	T-cell surface glycoprotein CD5	CD5
P10747	T-cell-specific surface glycoprotein CD28	CD28
P48643	T-complex protein 1 subunit epsilon	CCT5

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

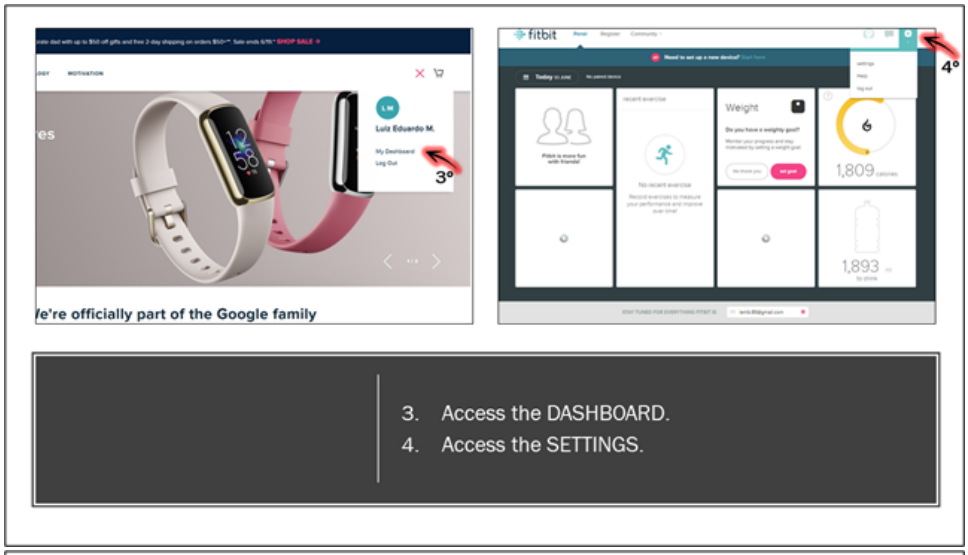
P37173	TGF-beta receptor type-2	TGFBR2
P51580	Thiopurine S-methyltransferase	TPMT
Q96J42	Thioredoxin domain-containing protein 15	TXNDC15
Q5JTD0	Tight junction-associated protein 1	TJAP1
P13726	Tissue factor	F3
P48307	Tissue factor pathway inhibitor 2	TFPI2
O43715	TP53-regulated inhibitor of apoptosis 1	TRIAP1
Q9Y6A5	Transforming acidic coiled-coil-containing protein 3	TACC3
P01375	Tumor necrosis factor	TNF
Q9NP84	Tumor necrosis factor receptor superfamily member 12A	TNFRSF12A
Q9NS68	Tumor necrosis factor receptor superfamily member 19	TNFRSF19
Q9HAV5	Tumor necrosis factor receptor superfamily member 27	EDA2R
P09758	Tumor-associated calcium signal transducer 2	TACSTD2
P00519	Tyrosine-protein kinase ABL1	ABL1
P07332	Tyrosine-protein kinase Fes/Fps	FES
P07948	Tyrosine-protein kinase Lyn	LYN
P07947	Tyrosine-protein kinase Yes	YES1
O15116	U6 snRNA-associated Sm-like protein LSm1	LSM1
Q9BSL1	Ubiquitin-associated domain-containing protein 1	UBAC1
Q8NBZ7	UDP-glucuronic acid decarboxylase 1	UXS1
P54727	UV excision repair protein RAD23 homolog B	RAD23B
Q8NEZ2	Vacuolar protein sorting-associated protein 37A	VPS37A
Q5VIR6	Vacuolar protein sorting-associated protein 53 homolog	VPS53
P49767	Vascular endothelial growth factor C	VEGFC
P17948	Vascular endothelial growth factor receptor 1	FLT1
P35968	Vascular endothelial growth factor receptor 2	KDR
P35916	Vascular endothelial growth factor receptor 3	FLT4
O95498	Vascular non-inflammatory molecule 2	VNN2
Q7Z5L0	Vitelline membrane outer layer protein 1 homolog	VMO1
Q6PCB0	von Willebrand factor A domain-containing protein 1	VWA1
Q96PQ0	VPS10 domain-containing receptor SorCS2 [Cleaved into: SorCS2 122 kDa chain; SorCS2 104 kDa chain; SorCS2 18 kDa chain]	SORCS2
Q7Z7D3	V-set domain-containing T-cell activation inhibitor 1	VTCN1
Q9Y5K8	V-type proton ATPase subunit D	ATP6V1D
Q8WWY7	WAP four-disulfide core domain protein 12	WFDC12
Q14508	WAP four-disulfide core domain protein 2	WFDC2
Q9Y5W5	Wnt inhibitory factor 1	WIF1
O43895	Xaa-Pro aminopeptidase 2	XPNPEP2
Q05516	Zinc finger and BTB domain-containing protein 16	ZBTB16
Q9UKS7	Zinc finger protein Helios	IKZF2

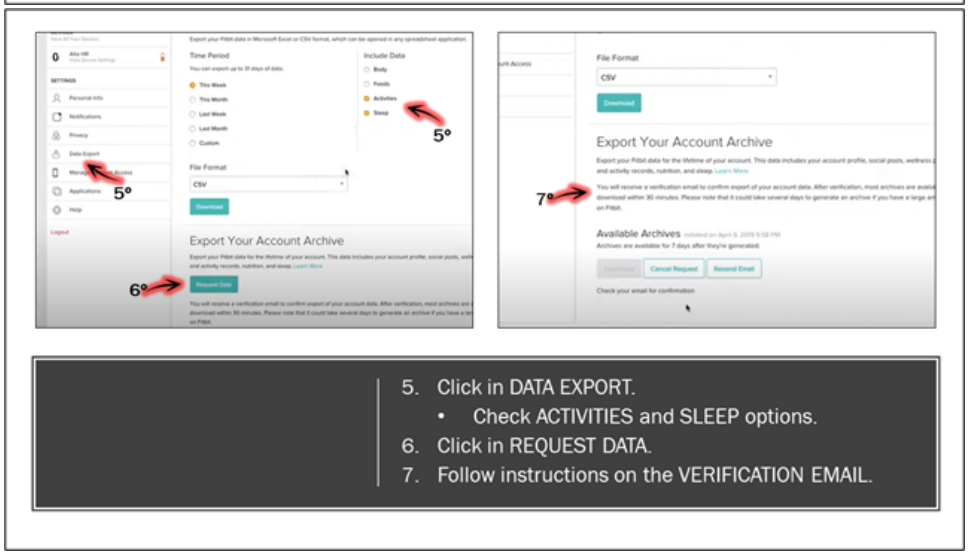
The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422

ANNEX VII : Wearable Cloud-based data extraction

The following is an example based on FitbitTM, given it has the largest market share and an established protocol for data sharing.

- Data is uploaded to Fitbit cloud services from the device.
- Users approve access to their data which may be accessed and shared to the MES-CoBraD Platform via API (see figures below)





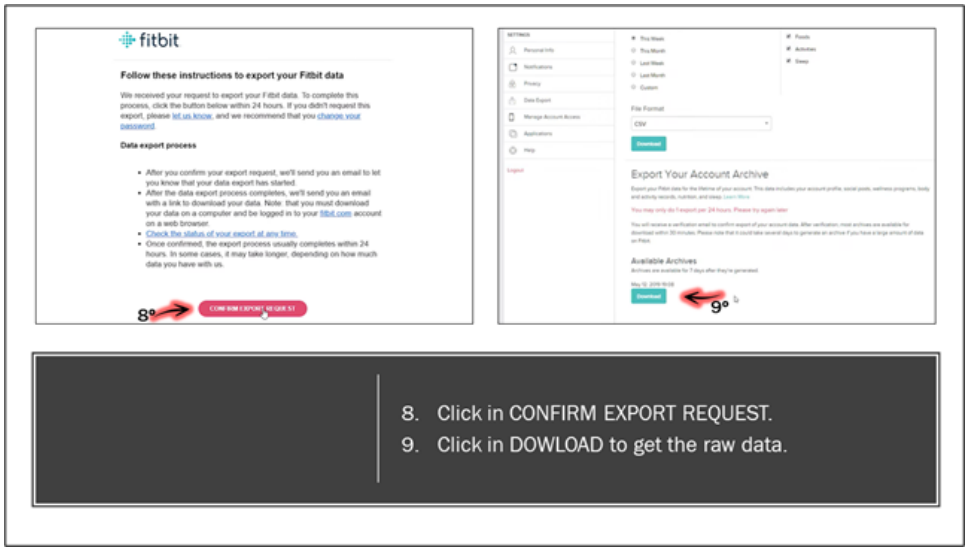


Figure 8: Fitbit User Access

After sharing is achieved, raw data (.json files) need to be converted, cleaned, and processed as part of the Advanced Analytics modules of the MES-CoBraD (these steps can be performed using Microsoft Power BI, or algorithms from open-source programming languages like R and Python).

The MES-CoBraD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 965422