D1.1 Data Management Plan

Deliverable No.	D1.1	**Due Date**	31/05/2023
Description	Quality and Data Management		
Type	Report	**Dissemination Level**	Sensitive
Work Package No.	WP1	**Work Package Title**	Project Management and Coordination
Version	V1.0	**Status**	Final

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement N° 101057747

Authors

Name and surname	Partner name	e-mail
Evita Karapintou Christoforos Papaioannou Chrysa Tsimperi Vasilis Tsakanikas Themis Exarchos Dimitrios Fotiadis	University of Ioannina	
Duccio Schiavon	Quantitas SA	
Doris Bamiou	University College London	
Dimitris Kikidis	National and Kapodistrian University of Athens	
Isabelle Walz	Universitaet Klinikum of Freiburg	
Noppawan Tundiew	King Chulalongkorn Memorial Hospital	

History

Date	Version	Change
01-04-2023	0.1	Table of Contents definition
10-05-2023	0.2	Deliverable draft for review
25-05-2023	0.3	Deliverable in the project template
31-05-2023	0.4	Deliverable revised and finalised
02-06-2023	1.0	Deliverable approved by the PC (ICCS) and submitted

Key data

Keywords	Prospective data, retrospective data, cohorts, registries, data use and re-use, data management plan
Lead Editor	Evita Karapintzou (UOI)
Internal Reviewer(s)	Konstantinos Bromis (ICCS), Doris Bamiou (UCL)

Abstract

The project consortium will establish a detailed data management plan as an annex to the Consortium Agreement. The TeleRehaB DSS data management plan outlines how data are to be handled both during the project, and after the project is completed. It addresses the many aspects of data management, metadata generation, data preservation, and analysis and ensures that data are well-managed in the present, and prepared for preservation in the future. The DMP describes the data management life cycle for the data to be collected, processed and/or generated by the TeleRehaB DSS project. All data will be available in digital form. The data collected and stored by following the FAIR data guideline principles.

The scope of this first version of the DMP is to answer all questions related to making research data findable, accessible, interoperable and re-usable (FAIR)[1] and provide information about TeleRehaB DSS compliance with FAIR principles. Research data should be 'FAIR', that is findable, accessible, interoperable and re-usable. These principles precede implementation choices and do not necessarily suggest any specific technology, standard, or implementation-solution. As part of FAIR data management, the TeleRehaB DSS will include information on:

- the handling of research data during and after the end of the project
- what data will be collected, processed and/or generated
- which methodology and standards will be applied
- whether data will be shared/made open access
- how data will be curated and preserved (including after the end of the project)

Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

[1] M. D. Wilkinson *et al.*, 'The FAIR Guiding Principles for scientific data management and stewardship', *Sci Data*, vol. 3, no. 1, Art. no. 1, Mar. 2016, doi: 10.1038/sdata.2016.18.

List of Abbreviations

Abbreviation	Explanation
AN	Acoustic Neuroma
BPPV	Benign Paroxysmal Positional Vertigo
BVW	Bilateral Vestibular Weakness
KCMH	King Chulalongkorn Memorial Hospital
CVD	Central Vestibular Dysfunction
DMP	Data Management Plan
FAIR	Findable, Accessible, Interoperable and Re-usable
MS	Multiple Sclerosis
MTBI	Mild Traumatic Brain Injury
NKUA	National Kapodistrian University of Athens
NPH	Normal Pressure Hydrocephalus patients
PD	Parkinson patients
PNP	Polyneuropathy patients
PPPD	Persistent Postural-Perceptual Dizziness
UCL	University College London
UKLFR	Universitaet Klinikum of Freiburg
UOI	University of Ioannina
UVW	Unilateral Vestibular Weakness
VM	Vestibular Migraine
WP	Work Package

Table of contents

About this deliverable

The TeleRehaB DSS concepts originate from the HOLOBALANCE and SMARTBEAR projects in which clinical trials assessed the feasibility of solutions for remote balance physiotherapy, supported by virtual AR couch. The TeleRehaB DSS consortium similarly includes partners that participate(d) in these two previous projects. The DMP is required for all projects participating in the extended Open Research Data (ORD) pilot, unless they opt out of the ORD pilot. However, TeleRehaB DSS is submitting a DMP to cover any case. The decision to or not to opt out is still being discussed within the project. Opting out is possible at any stage (even after signing the grant) and in TeleRehaB DSS the two main issues that need further exploration before participation in the ORD is finalized are:

- Participation may be incompatible with the obligation to protect results that can reasonably be expected to be commercially or industrially exploited. This can only be defined by the end of year 2 and final decisions may have to be made by the end of the project when exploitation aspects will be finalized and agreed among the partners.

- Participation may be incompatible with rules on protecting personal data, especially as in TeleRehaB DSS raw and processed data is foreseen. This incompatibility will be explored along with ethical issues and protocol approvals since sharing the data openly should be clearly mentioned in the information sheets and consent forms according also with GDPR and any reuse of or open access to data should be made perfectly clear to the consenting patients.

The TeleRehaB DSS research data will be 'FAIR', that is findable, accessible, interoperable and re-usable and the approach to achieve that are described in this deliverable.

The deliverable is structured according to the Horizon 2020 FAIR Data Management Plan.

1 Data Summary

1.1 Purpose of the data collection/generation and its relation to the objectives of the project

Data collection is foreseen in WP5.

Population to be studied and sample

Patients aged 50-80 years old with falls/at risk of falls AND/OR with chronic dizziness/imbalance AND with radiologically diagnosed stroke, OR diagnosis of mild cognitive impairment (MCI) OR peripheral or mixed (central and peripheral) vestibular diagnosis OR long Covid-19. The clinical study will be conducted from month 12 to month 32. TeleRehaB DSS will recruit equal numbers of the same diagnostic group. Also, it will actively recruit patients who are users of wearables & IoT devices with high tech literacy skills, who may bring their own data to the study.

The sample size for the pilot study will be 50 participants for each intervention group due to the number of variables included, with 50 stroke patients, 50 MCI, 50 vestibular, 50 long Covid in the intervention arm and equal number of age- sex matched controls of the same diagnosis per number of subjects recruited in each centre to account for country specific factors.

	Site	Targeted patient population	Number of patients
Study 1	UCL (United Kingdom)	Stroke, MCI, long Covid-19	50 intervention + 50 controls
Study 2	NKUA (Greece)	Vestibular, MCI	50 intervention + 50 controls
Study 3	UKLFR (Germany)	Stroke, MCI, long Covid-19	50 intervention + 50 controls
Study 4	KCMH (Thailand)	Stroke, MCI, Vestibular	80 intervention + 40 controls
Study 5	RYDP (Spain)	Stroke, Vestibular, long Covid-19	50 intervention + 50 controls
Overall:			280 intervention + 240 controls

1.2 Types and formats of data the project generated/collected

Randomised Controlled Trial study – conducted for clinical validation purposes

During the TeleRehaB DSS a Clinical Validation Study led by UCL will be conducted in the 5 clinical partner settings i.e. UCL (United Kingdom), UKLFR (Germany), NKUA (Greece), KCMH (Thailand), RYDP (Spain) (see table in 2.1). The following data will be collected:

Personal data These will include name, date of birth, address, contact details (email, phone number), General Practitioner doctor -GP name and contact details and will be kept for the duration of the study in the Data Safe Havens of each clinical partner. These data are necessary in order to contact the participants throughout the course of the study, and their GP to notify of the participants' involvement in the study and any significant findings during it. The research team will also collect special category data – relating to the participants' health.

In addition to the personal data the following data will be collected and analysed:

1. Baseline measures will include demographic data (age, occupation, years of education, sex, married status) and all primary and secondary outcome measures (see below).

2. Primary outcome measures:

a. We will follow previously published methodology to calculate economic costs of the interventions. For the control arm, these will be the printing costs (if any) of the booklets; employer and overhead costs per minute for each therapist increased by a multiplier calculated from the ratio of client contact time to total time obtained from a published source of unit costs patient utilization of primary care and hospital services.

b. We will calculate quality-adjusted life years (QALYs) with the EuroQol five dimensional descriptive system EQ-5D-5Linstrument (https://euroqol.org/eq-5d-instruments/) and QALY weights derived using cross-mapping to the EQ-5D-3L UK values and other European tariffs, as recommended by NICE and with an area-under-the curve approach, assuming linear change in EQ-5D-5L values over time with primary end point at the end of intervention. We will establish cost-effectiveness by means of the incremental cost-effectiveness ratio (ICER), that is the difference in mean costs divided by the difference in mean QALYs at time of recruitment to the study and 1 year before the study and for 1 year after the end of intervention.

3. Secondary outcome measures

Secondary outcomes will include the same measures as per our published HOLOBalance protocol (https://clinicaltrials.gov/ct2/show/NCT04053829) to be collected at baseline (week 0) and end of intervention follow up (week 12). These will include functional measures of dizziness (such as Dizziness Handicap Inventory); balance and gait function (functional gait assessment), balance confidence (ABC Scale), fear of falling (FES-I) and prospective falls; cognitive impairment (MoCA); mood (hospital anxiety depression scale HADS); treatment adherence (patient exercise log diary); real life sensor recorded physical activity.

a. Balance Assessment measures:

- The mini Balance Evaluation Systems Test BESTest (a 14-item test that assesses dynamic balance with a total score of 28 points, 10 minutes). (Reference:

Franchignoni, F., et al., Using psychometric techniques to improve the Balance Evaluation Systems Test: the mini-BESTest. J Rehabil Med, 2010. 42(4): p. 323-31). From: https://www.bestest.us/files/7413/6380/7277/MiniBEST_revised_final_3_8_13.pdf

- The 10-item Functional Gait Assessment (FGA) test that assesses complex gait tasks (e.g. walking with head turns or stopping and turning, 5 minutes). (Reference: Wrisley DM, Kumar NA. Functional gait assessment: concurrent, discriminative, and predictive validity in community-dwelling older adults. Phys Ther 2010;90:761–73). From: https://geriatrictoolkit.missouri.edu/FGA/Wrisley-2007-FGA_PTJ_84-10-Appendix.pdf

- Falls diaries will be collected monthly for the duration of the intervention and for 3 up to 6 months after completion of the intervention.

b. Cognitive Assessment measures:

The validated Montreal Cognitive Assessment (MoCA) that includes sections on visuospatial/executive function, naming, attention, language, abstraction, memory and orientation to time and place (6 questions) (15 minutes). (Reference: Nasreddine, Z.S., et al., The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc, 2005. 53(4): p. 695-9). From: https://www.parkinsons.va.gov/resources/MOCA-Test-English.pdf

c. Physical Activity and Social Participation Assessment:

The 9-item self-administered Rapid Assessment of Physical Activity (RAPA) is a questionnaire that assesses levels of a wide range of physical activity level in adults older than 50 years (5 minutes). (Reference: Topolski TD, LoGerfo J, Patrick DL, et al. The rapid assessment of physical activity (RAPA) among older adults. Prev Chronic Dis 2006;3:A118). From: http://depts.washington.edu/hprc/programs-tools/tools-guides/rapa/

d. Subjective Questionnaires

- The 25-item self-report Dizziness Handicap inventory - DHI validated questionnaire that assesses functional, emotional and physical domains. Responses are graded 0 (no), 2 (sometimes) or 4 (yes) with higher scores indicating greater impact of dizziness maximum and maximum score of 100 points (15 minutes). (Reference: Jacobson, G.P. & C.W. Newman, The development of the Dizziness Handicap Inventory. Arch Otolaryngol Head Neck Surg, 1990. 116(4): p. 424-7). From: https://geriatrictoolkit.missouri.edu/vest/Dizziness-Handicap-Inventory.pdf

- The Activities-specific Balance Confidence Scale (ABC) that assesses patient's perceived confidence for 16-activities of daily living without losing balance. Scores ≤67/100% indicate increased falls risk (10 minutes). (Reference: Powell, L.E. and A.M. Myers, The Activities-specific Balance Confidence (ABC) Scale. J Gerontol A Biol Sci Med Sci, 1995. 50A(1): p. M28-34). From: https://sites.temple.edu/rtassessment/files/2018/10/Activities-Specific-Balance-Confidence-ABC-Scale.pdf

- The Falls Efficacy Scale International (FES-I) that measures an individual's level of concern regarding falling during social and physical activities inside and outside the

home is measured on a four-point Likert scale (1=not at all to 4=very; 10 minutes). (Reference: Delbaere, K., et al., The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study. Age and Ageing, 2010. 39(2): p. 210-216). From: https://documents.manchester.ac.uk/display.aspx?DocID=38565

- The EQ-5D-5L, a standardized, valid and reliable simple, generic measure of health status for clinical and economic appraisal. The respondent is asked to rate their health status on these five dimensions from 1 to 5 respectively as no problems, slight problems, moderate problems, severe problems, and extreme problems. The EQ VAS (Visual Analogue Scale) records the respondent's self-rated health on a 20 cm vertical, visual analogue scale with endpoints labelled 'the best health you can imagine' and 'the worst health you can imagine'. The respondent is asked to mark an X on the scale to indicate "how your health is TODAY". (Reference: EuroQol - a new facility for the measurement of health-related quality of life. Health Policy, 1990. 16(3): p. 199-208). From: https://euroqol.org/eq-5d-instruments/sample-demo/

- The Hospital Anxiety and Depression Scale (HADS), a 14-item scale which assesses non-somatic anxiety (HAD-A) and depression (HAD-D) symptoms. Scores range from 0-21 for each subscale and a score ≥8 identifies depression and anxiety. (References: Zigmond, A.S. & R.P. Snaith, The hospital anxiety and depression scale. Acta psychiatrica scandinavica 1983; 67(6): 361-370. Bjelland I, Dahl AA & Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res 2002; 52(2): 69-77). From: https://www.svri.org/sites/default/files/attachments/2016-01-13/HADS.pdf

e. We will also collect treatment adherence (patient exercise log diary); and real life sensor recorded physical activity (from patients who already track this activity via their smart phones and watches).

f. We will also assess usability of the TeleRehaB DSS system by collecting the following two questionnaires from research team members and participants enrolled in the intervention group only:

- The system Usability Scale (SUS) (collected at week 12 only). From: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html

- User Experience Questionnaire (UEQ) (Collected at weeks 0,12 and 8). From: https://www.ueq-online.org/

g. Exit interviews will be held to collect qualitative data to explore participants' experience of the exercise programme and will be recorded, transcribed verbatim and subjected to thematic analysis to determine commonalities in their experiences.

Proposed Analysis Plan for RCT Data

Data will be presented as mean ± SD. The chi-square test will be used to assess how drop-out rate and demographic data between study groups. Between-group pre-post intervention analysis (IG vs. CG) will be conducted using a general multivariate linear model (MANOVA). Assumption testing applied correlation and chi-square

analyses will be conducted as appropriate. We will employ intention-to-treat principles, and all participants randomised to an intervention will be included in analysis. MANOVA Statistical adjustment will be conducted if group differences in baseline dependent variables were statistically significant. Alpha level will be set at p= 0.05.

Responses to exit interview questions will be analysed using a Dovetail application (https://dovetail.com) which facilitates the analysis and management of user research data by providing qualitative data analysis tools. For each question, participant responses were manually tagged and a chart was created using the application. As a result, responses were reported qualitatively as numbers, with inclusion of some verbatim exemplar statements.

1.3 Re-use of existing data in TeleRehaB DSS and their origin

The pre-clinical evaluations will be based on developing the tools that are necessary to assess the patient's current status and most importantly evaluate his/her risks for imminent deterioration in his/her well-being over the near future and the risk for occurrence of detrimental incidents such as falls. Retrospective data will be generated and collected for module harmonization. The retrospective data will be questionnaires and physical activity data. The collected data will contribute to the extraction modules of prognostic and risk factors, fall risk factors, treatment effectiveness, side effects complications and compliance factors.

A rich volume of retrospective patient data from previous projects, clinical partner datasets and open EU databases will be used to identify the multimorbidity key aspects that are related to balance/gait deficiencies, reduced independence and level of activity (ADL, moving outdoors, social activities, etc.), and falls, by examining a wide range of medical data and patient history. In totals, the retrospective data from the pilot partners are from NKUA (250), UKLFR (214), KCMH (9200), UCL (414). Secure data transfer is ensured with agreement for the sharing of anonymised data with TeleRehaB DSS consortium in relation to the research project (A.1.1 Appendix).

Source of retrospective data	Subjects
HOLOBALANCE	160
NKUA (Greece)	250
UKLFR (Germany)	214
KCMH (Thailand)	9,200
UCL (UK)	414

UK Biobank (UCL)	500,000
Total	510,238

1.3.1 Types and formats of retrospective data

Previous project data

The dataset from HOLOBALANCE concerns at total 80 intervention and 80 control group, aged >40 years old. The clinical evaluation data and sensor data include:

- Demographics, clinical evaluations and questionnaires such as EQ-5D5L, MoCA, SUS, FGA, ABC, RAPA, MiniBest, cognitive function
- Exercise data from 2 IMUs, pressure insoles, chest HR monitor, depth sensor
- Raw data from smartwatch (Steps, activity capture)

OTHER RETROSPECTIVE STUDIES

Study by NKUA (GREECE)

- Study Design

This analysis will look for the association between falls risk and various factors in patients with MCI and vestibular disturbances.

- Study population

The retrospective data from NKUA concerns at total 250 subjects and collecting demographics, symptoms, severity of symptoms, clinical signs, comorbidities, test for balance, falls. The data from NKUA concerns a of total 250 subjects from 2 different datasets. The first dataset corresponds to 104 subjects (29 PPPD patients, 20 VM patients, 15 UVW patients, 19 Meniere's disease patients, 3 AN patients, 7 BPPV patients, 5 BVW patients, 2 CVD patients, 1 mal de debarquement patient, 1 MS patient, 2 MTBI patients) and the second dataset to 146 subjects (74 PPPD patients, 26 VM patients, 24 BPPV patients, 8 Meneire's disease patients, 7 Neuritis patients, 3 UVW patients, 1 CVD patients, 3 fallers).

The 1st dataset will be:

- Demographic and Clinical Information (age, gender, diagnosis, comorbidities, clinical signs).
- Scales: DHI, FGA (2 time-points: baseline, final rehabilitation session).
- Tests: Tandem Romber, Unterberger.

The 2nd dataset will be:

- Demographic and Clinical Information (age, gender, diagnosis, education level).

- Scales: Dizziness Handicap Inventory (DHI), Functional Gait Assessment (FGA), Activities Specific Balance Confidence Scale (ABC) (2 time-points: baseline, final rehabilitation session), Montreal Cognitive Assessment (MoCA).

Study by KCMH (Thailand)

- Study design

This retrospective data analysis aims to investigate the association between fall risk and various factors among older adults in Thailand. The study will use existing data on older adults' medical history, demographic information, medication use, and fall history to identify potential risk factors for falls. The anonymous data will be analyzed in partnership with EU partners.

- Study population

The retrospective study will involve reviewing the medical records of the 9,200 older adults and collecting demographics, cognitive function, depression screening, falls history, sarcopenia test, urinary incontinence, polypharmacy.

Inclusion criteria for retrospective research participants:

1. Age: 60 years or older
2. Availability of medical records
3. Living in Thailand

Exclusion criteria for retrospective research participants:

1. Age: Below 60 years old
2. Incomplete or missing medical records.
3. Living outside of Thailand

Information on demographic factors such as gender, age, race, ethnicity, and body mass index will be gathered from the patients' medical records. The researchers will also extract additional data from the charts regarding the patients' fall history, ability to perform daily activities (Barthel Index, The Lawton Instrumental Activities of Daily Living (IADL) Scale, ADL), level of dependency, depression, and cognitive function (MoCA), nutritional status (MNA-Short and Full), insomnia, nocturia, grip strength, gait speed, and the presence of sarcopenia. Additionally, the researchers will collect dates related to the use of multiple medications, particularly sleep pills.

Study by UKLFR (Germany)

The data from UKLFR concerns at total 214 participants (60 healthy, 20 PNP patients, 35 PD, 28 L-dopa tests, 21 NPH patients, 50 Holobalance subjects. The data will be:

- Demographic and Clinical Information (age, gender, weight, BMI, diagnosis, comorbidities, clinical signs, medications).
- Raw data from 12 cameras (Full body motion capture- gait speed, step length, step time, cadence, double support phase).

Study by UCL (United Kingdom)

The dataset from UCL concerns at total 414 subjects from 5 different datasets. The 1st dataset corresponds to 56 participants and is a comparison of older adults who recruited from neuro-otology (15), falls clinic for balance assessment (25) and 16 healthy controls. The data will be:

- Demographics and Clinical Information (age, gender, clinical signs- fall in past year).
- Scales: Functional gait assessment (FGA), Computerized dynamic posturography (CDP), Risk Physiological Profile Assessment (PPA), Vertigo Symptom Scale – vestibular subscale (VSS-V), Vertigo Symptom Scale – anxiety subscale (VSS-A), Vestibular Disorders Activities Daily Living –Functional, -Ambulation, -Instrumental (VDADL-F,-A,-I), Hospital Anxiety and depression scale – Depression subscale, – Anxiety subscale (HAD-D,-A), Scale Activities Specific Balance Confidence scale (ABC).
- Questionnaires: Situational characteristic Questionnaire (SCQ).

The 2nd UCL dataset corresponds to 17 participants and is data from a study comparing OTAGO supplemented with MSR exercises. Data will be the same as 1st with the difference that it is collected at baseline (variable_0), 4 weeks (variable_4) and 8 weeks (variable_8).

The 3rd UCL dataset corresponds to 104 participants and will be:

- Demographics and Clinical Information (age, gender, BMI, clinical signs)
- Scales: Vertigo Symptom Scale – vestibular subscale (VSS-V), Dizziness Handicap Inventory (DHI), Scale Activities Specific Balance Confidence scale (ABC), Hospital Anxiety and depression scale (HAD), Functional Gait Assessment (FGA), Physical Activity Level (PAL), Vertigo Symptom Scale (VSS)
- Questionnaires: Situational characteristic Questionnaire (SCQ).

The 4th UCL dataset corresponds to 153 subjects. 93 are a group of patients with vestibular migraine, with or without TBI, and 60 are with TBI (Traumatic Brain Injury). The data will be:

- Demographics and Clinical Information (age, gender, diagnosis, clinical signs)
- Scales: Dizziness Handicap Inventory (DHI), Visual Analog Scale (VAS), Functional Gait Assessment (FGA) (In 2 time-points: pre and post physiotherapy).
- Questionnaires: Sexual function-Vaginal changes Questionnaire (SVQ) (In 2 time-points: pre and post physiotherapy).
- Tests: Modified Clinical Test of Sensory Interaction in Balance (mCTSIB), Dynamic Visual Activity (DVA) (In 2 time-points: pre and post physiotherapy).

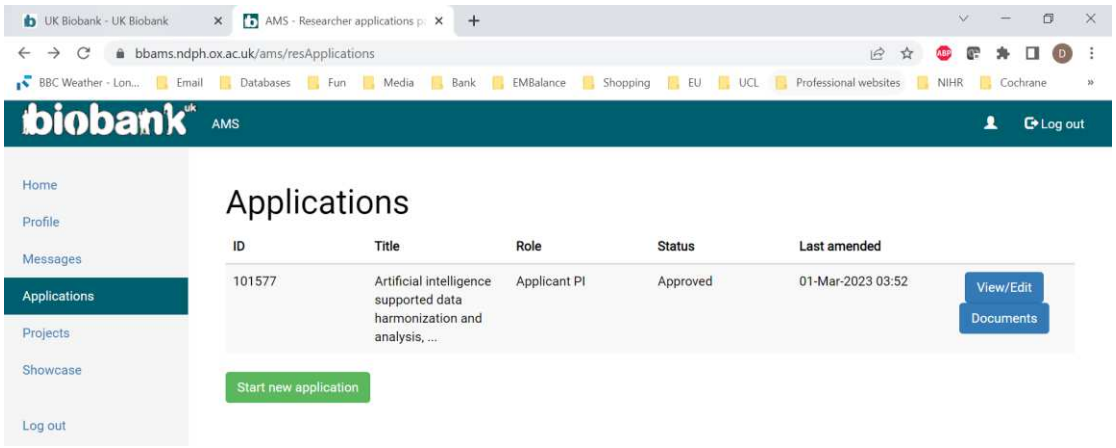
The 5th UCL dataset concerns to 84 participants from 2 different studies. The first corresponds to 51 subjects (30 healthy and 21 with unilateral peripheral vestibular disorder) and the other to 33 subjects (16 healthy and 17 stroke). The data will be:

- Demographic and Clinical Information (age, gender diagnosis, clinical signs, symptoms duration).
- Scales: Functional Gait Assessment (FGA), Hospital Anxiety and depression scale (HAD), Scale Activities Specific Balance Confidence scale (ABC), Functions of Music Listening scale (FMUL).
- Tests: Time up and go test (TUG)
- Raw data (steps, walk duration, velocity).

UCL -UK Biobank data

In addition to the above retrospective data, UCL has applied to the UK Biobank https://www.ukbiobank.ac.uk/ to purchase and obtain access for analysis purposes to their existing dataset. UK Biobank is a large-scale biomedical database and research resource, containing in-depth genetic and health information from half a million UK participants. The database is regularly augmented with additional data and is globally accessible to approved researchers undertaking vital research into the most common and life-threatening diseases. It is a major contributor to the advancement of modern medicine and treatment and has enabled several scientific discoveries that improve human health.

The application was approved (see screenshot below) and the material transfer agreement -MTA is provided in the annex.



The Biobank application was entitled "Large scale retrospective data harmonization and analysis using AI, for the identification of prognostic factors related to balance/gait disorders and risk of falls".

We proposed to conduct analysis of Biobank patient data to identify subprofiles and clustering of multimorbidity factors that predict falls after correcting for age socioeconomic and other factors.

Aims: To identify behavioral (incl. physical activity measured by 24-hour physical activity monitor), psychological/psychiatric, cognitive, neurological (incl. Brain imaging) and medical factors that predict falls.

The primary outcomes will be to: a) evaluate the effectiveness of the treatments used in each patient, isolate cases depending on the level of response to treatment depending on their diagnosis group, focusing on cases of positive and negative responses as well unresponsive subjects, b) isolate cases where abnormal adverse and side effects have been provoked as a result of the suggested treatment plan and evaluate their baseline characteristics for prognostic factors, and c) analyse the cost-

effectiveness of each intervention path, taking into consideration the extra burden of dealing with provoked side effects in terms of human effort and institutional costs (e.g. prolonged treatment, in clinic visits, additional medications, etc.).

Background and scientific rationale of the proposed research project: There is a lack of understanding of the potentially synergistic effect of the different constellations of coexisting chronic conditions, together with psychosocial and behavioral factors of the individual, on the impact that they have on the affected individual. This impact goes beyond a simple count of the conditions, or even their weighted impact. It has been proposed that elucidating disease specific endotypes, i.e. the biology of how different chronic conditions interact and impact on a given condition will help identify endotype-directed interventions that may ameliorate clinical outcomes. We will conduct analysis of multimorbidities from Biobank to identify subprofiles and clustering of multimorbidity factors. A discovery based objective approach should allow an evaluation of the respective contribution of different risk factors in different systems which will then drive hypothesis driven therapeutic approaches to patient management.

Brief description of the methods: The two basic methodologies to be used are data harmonization and data analytics using AI. Data harmonization aims to overcome the structural heterogeneities that are present among the medical data from multiple sources and transform them into a common format with the same parameters and range values, using data driven and computational approaches. Two different approaches will be used to accomplish data harmonization, namely the stringent and the flexible strategy, depending on the specifics of each dataset merge. The former limits the harmonization process only on data that will be (or have been) collected under common measurement procedures (standards) whereas the latter approach extends the harmonization process to include data that have been already collected under different measurement procedures or protocols (which is the case for UK Biobank data and their harmonization with the other sources). The majority of the methods for retrospective data harmonization make use of a pre-defined, standardized model which describes the requirements of a particular clinical domain and serves as a common template (i.e., a gold standard) for harmonization. A robust data harmonization method involves the application of lexical and semantic matching algorithms. A lexical matching algorithm uses string similarity techniques to identify common terminologies (i.e., exact sequences or similar block sequences) that are present between the terms of the standard model and those from the original datasets. Semantic matching uses a standard (or reference) model which is usually expressed in the form of an ontology, where the classes are considered as categories. The outcome of the data harmonization engine will be fed to the AI models for subsequent analysis and evaluations. AI models for prognostic analytics AI data-driven models are relatively less common as most of the evidence in the field comes from studies using regression models and statistical analysis, which might not be the most optimal way to ensure holistic prognostic factors if not considering multifactorial analytics when performing hypothesis testing in balance deficits. The algorithms mine vast amounts of complex data (e.g. clinical characteristics and disease expression profiles), looking for combinations and patterns that determine response to a specific biologic agent. AI-based machine learning algorithms offer the potential to conduct prognostic hypothesis testing of future events (e.g. falls, adverse events, treatment side effects), and optimize personalized treatment options and subsequently provide clinicians with actionable models for making better treatment decisions and allowing for a more (cost-)effective disease management. Falls predictors will be isolated using machine learning techniques and compared to statistical analysis and previous findings. Among the proposed algorithms to be evaluated to formulate the predictive

model are regression models, simple classifiers (e.g. Naïve Bayes, decision trees, rule-based classification, etc.) and more advanced techniques involving Random Forest and boosting algorithms. Where longitudinal data streams are available, methods such as tree-based ensemble predictive models, trajectory-based analysis, latent class modelling (LCM) and Dynamic Bayesian Networks (DBNs), while also expanding to deep learning techniques and recursive networks such as Long Short-Term Memory (LSTM) models.

Type and size of dataset required: We will include all participants who gave a positive answer (subjects) and a negative answer (controls) to the following question "In the last year have you had any falls?" and we will document the number of falls reported by positive respondents.

We will assess the following type of data:

Medical information (Eye problems/disorders, vascular/heart problems diagnosed by doctor, long-standing illness, disability or infirmity, other serious medical condition/disability, hearing difficulty/problems diagnosed by doctor medication for cholesterol, blood pressure or diabetes, participants BMI, height, weight, pulse, overall health rating,

Physical: impedance of leg, overall acceleration summary (check other:https://biobank.ndph.ox.ac.uk/showcase/label.cgi?id=1006)

genetic principal components

Socioeconomic: Employment status, Education/professional qualifications

Biomarkers

Cognitive: Fluid intelligence Pairs matching, Prospective memory, reaction time

Behavioural: Typical diet; Alcohol consumption; Physical activity assessed by a 24-hour activity monitor for a week (incl repeated measures in some)

Psychological/Psychiatric/Mental health: Self reports (Ever felt worried, tense, or anxious for most of a month or longer; Recent feelings or nervousness or anxiety; Recent feelings of depression; Recent feelings of tiredness or low energy; etc.)

Imaging data: T1 structural brain MRI; Resting functional brain MRI; Total volume of white matter hyperintensities

Expected value of the research: The obtained data from the UK biobank will be analysed in order to define prognostic factors that contribute to the risk of falls and balance disorders. The retrospective data analytics consists of four modules, each one of which is targeting a specific factor including evaluations for prognosis of risk factors related to:

a. general well-being deterioration and risk of falls
b. response to treatment
c. treatment side effects
d. compliance and adherence to use of novel technologies

Apart from the UK Biobank data, a rich volume of retrospective patient data from previous projects, clinical partner datasets and open EU databases will be used to identify the multimorbidity key aspects that are related to balance/gait deficiencies,

reduced independence and level of activity (ADL, moving outdoors, social activities, etc.), and falls, by examining a wide range of medical data, patient history, cognition, genetics and imaging scans. A secondary analysis will be conducted to identify the counteracting factors that may reduce risk of falls and empower individuals to be more active. Factors that define response to rehabilitation treatment and prognosis of potential adverse/side effects of treatment are particularly difficult to analyse and model effectively since there is limited access to open data of systematic clinical trials. We will capitalise on the unique opportunity to analyse diverse granular clinical data from different previous projects, and patient clinical records of the clinical partners to form and analyse a comprehensive database of intervention outcomes utilising of standard of care, AR systems, gamified rehabilitation exercises and cognitive training.

PLEASE NOTE THAT THE UK BIOBANK DATA WILL BE MADE ACCESSIBLE TO UOI FOR ANALYSIS WITHIN THE UCL DATA SAFE HAVEN ONLY FOR DATA GOVERNANCE PURPOSES AND ACCORDING TO THE LEGALLY BINDING MTA AGREEMENT SIGNED BY UCL. THESE DATA WILL NOT BE MADE AVAILABLE TO OTHER RESEARCHERS BY THE CONSORTIUM AS THEY ARE OWNED AND GOVERNED BY BIOBANK.

1.4 To whom might the datasets be useful ("data utility")

Researchers studying:

- Vestibular disorder
- Long COVID-19
- Mild Cognitive Impairment
- Stroke
- Neurologists
- ENTs
- Human balance and its rehabilitation
- Human-machine interactions
- Signal processing
- Medical data mining

The datasets generated by the TeleRehaB DSS project are likely to be of substantial relevance to a wide variety of researchers and practitioners. These databases will give rich, multidimensional insights into numerous aspects of human health and rehabilitation, making them a valuable resource for a wide range of studies and applications.

Researchers researching vestibular diseases, long COVID-19, moderate cognitive impairment, and stroke can benefit from these datasets since they provide a wealth of information regarding the impact of these ailments on patients' balance and mobility. This information can aid in the formulation of more efficient treatment and

rehabilitation plans as well as aid in understanding how these disorders advance and how they affect patients' quality of life.

Similarly, this information can be used by neurologists and ENTs to learn more about the neurological and physiological underpinnings of balance and movement issues. This can support the identification and management of various illnesses and support the advancement of novel therapy modalities.

These datasets can also be used by researchers working in the areas of human balance and its rehabilitation, human-machine interfaces, signal processing, and medical data mining. The data can provide information about how well different rehabilitation tactics work, how people and rehabilitation devices interact, and patterns and trends in the data that might guide future study and treatments.

2. FAIR data

The architecture and the data storage will take into account the FAIR principles, and will work towards adopting EOSC Interoperability Framework to allow the technical, semantic, organizational and legal interoperability. The FAIR principles are a set of guidelines for making data more discoverable, accessible, interoperable, and reusable. FAIR Principles in TeleRehaB DSS are ensured by selecting Zenodo (zenodo.org) as the repository for depositing both publications and data and by adopting the DataCite Metadata Schema (schema.datacite.org).

2.1 Making data findable, including provisions for metadata

- **F1**. (Meta)data are assigned a globally unique and persistent identifier
    - A persistent and unique Digital Object Identifier (DOI) is issued to every published record on Zenodo. Moreover, DOI versioning is supported and enables users to update the record's files after they have been made public and researchers to easily cite either specific versions of a record or to cite, via a top-level DOI, all the versions of a record.
- **F2**. Data are described with rich metadata (defined by R1 below)
    - Zenodo's metadata is compliant with DataCite's Metadata Schema minimum and recommended terms, with a few additional enrichments. As there are no specific metadata schemas that can be used with the TeleRehaB DSS data this more generic schema will be adopted.
- **F3**. Metadata clearly and explicitly include the identifier of the data it describes
    - The DOI is a top-level and a mandatory field in the metadata of each record.
- **F4**. (Meta)data are registered or indexed in a searchable resource

- Metadata of each record is indexed and searchable directly in Zenodo's search engine immediately after publishing.
- Metadata of each record is sent to DataCite servers during DOI registration and indexed there.

2.2 Making data openly accessible

- **A1**. (Meta)data are retrievable by their identifier using a standardized communications protocol
  - Metadata for individual records as well as record collections are harvestable using the OAI-PMH protocol by the record identifier and the collection name.
  - Metadata is also retrievable through the public REST API.
- **A1.1**. The protocol is open, free, and universally implementable
  - See point A1. OAI-PMH and REST are open, free and universal protocols for information retrieval on the web.
- **A1.2**. The protocol allows for an authentication and authorization procedure, where necessary
  - Metadata are publicly accessible and licensed under public domain. No authorization is ever necessary to retrieve it.
- **A2**. Metadata are accessible, even when the data are no longer available
  - Data and metadata will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least.
  - Metadata are stored in high-availability database servers at CERN, which are separate to the data itself.

2.2.1 Participation in ORD

The two main issues that need further exploration before participation in the ORD is finalized are:

- Participation may be incompatible with the obligation to protect results that can reasonably be expected to be commercially or industrially exploited. This can only be defined by the end of year 2 and final decisions may have to be made by the end of the project when exploitation aspects will be finalized and agreed among the partners.
- Participation may be incompatible with rules on protecting personal data especially since in TeleRehaB DSS raw and processed data from older citizens in home balance rehabilitation is foreseen. This incompatibility will be explored along with ethical issues and protocol approvals since sharing the data openly should be clearly mentioned in the information sheets and consent forms according also with GDPR and any reuse of or open access to data should be made perfectly clear to the consenting patients.

2.2.2 Methods and/or software tools needed to access the data

Approval from the TeleRehaB DSS Board for accessing patient data will be requested.

Viewing the data is possible through commonly used software such as video viewers and text editing software.

For processing the data Matlab or any similar software is needed. Two open source alternatives to Matlab could be:

1. **GNU Octave** (www.gnu.org/software/octave/)
2. **Scilab** (www.scilab.org)

For using the data within the TeleRehaB DSS concept the system developed in the project is needed.

2.2.3 Zenodo Repository

Zenodo is built and developed by researchers, to ensure that everyone can join in Open Science. The OpenAIRE project, in the vanguard of the open access and open data movements in Europe was commissioned by the EC to support their nascent Open Data policy by providing a catch-all repository for EC funded research. CERN, an OpenAIRE partner and pioneer in open source, open access and open data, provided this capability and Zenodo was launched in May 2013. In support of its research programme CERN has developed tools for Big Data management and extended Digital Library capabilities for Open Data. Through Zenodo these Big Science tools could be effectively shared with the long--tail of research.

Zenodo features and policies cover TeleRehaB DSS needs. Specifically:

- It supports DOI versioning ZENODO ensures the discovery and citability of the research output by assigning a Digital Object Identifier (DOI) to every upload
- It supports Flexible licensing
- It is integrated with GitHub which promotes software citation and preservation through one-click
- They currently accept up to 50GB per dataset (one can have multiple datasets). There is no size limit on communities.
- All research outputs from all fields of science can be stored. In the upload form you can choose between types of files: publications (book, book section, conference paper, journal article, patent, preprint, report, thesis, technical note, working paper, etc.), posters, presentations, datasets, images (figures, plots, drawings, diagrams, photos), software, videos/audio. i.e. all the types of data in TeleRehaB DSS
- The data is stored in CERN Data Center. Both data files and metadata are kept in multiple online and independent replicas. CERN has considerable knowledge and experience in building and operating large scale digital repositories and a commitment to maintain this data centre to collect and store 100s of PBs of LHC data as it grows over the next 20 years. In the highly unlikely event that Zenodo will have to close operations, they guarantee that they will migrate all content to other suitable repositories, and since all uploads have DOIs, all citations and links to Zenodo resources (including TeleRehaB DSS data) will not be affected.

2.2.4 Restrictions on use – Data Management Board

In the pilot data from patients will be generated and collected. Thus, the intended data use and the purpose of the studies needs to be clearly defined.

A data management board will be constituted and will decide about controlled access and relevant conditions.

The board consists from one representative from each organization participating in the pilot and the project coordinator. The members are:

- Prof. D. I. Fotiadis from UOI
- Prof. G. Matsopoulos from ICCS
- Dr. N. Utoomprurkporn from KCMH
- Mr. A. Vontas from VILABS
- Prof. N. Filipovic from BIOIRC
- Dr. S. Guillen from ACT
- Prof. A. Bibas from NKUA
- Prof. D. Bamiou from UCL
- Prof. G. Cea from BRD
- Prof. R. Goncalves from NOVA
- Mr. D. Schiavon from QUANE
- Prof. C. Maurer from UKLFR
- Ms. K. Richards from HIN
- Dr. S. Ortuño from RYDP

2.2.5 Well described conditions for access (i.e. a machine readable license)

A Creative Commons Attribution-Non Commercial ShareAlike (CC-BY-NC-SA) license seems to be the one that fits for TeleRehaB DSS datasets, and especially for the dataset generated within the pilot.

Restricted/ controlled access means that only authenticated and authorized users whose research proposal has been vetted by the data management Board will be provided access to the pilot dataset.

2.3 Making data interoperable

- **I1**. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
    - Zenodo uses JSON Schema as internal representation of metadata and offers export to other popular formats such as Dublin Core or MARCXML.
- **I2**. (Meta)data use vocabularies that follow FAIR principles
    - For certain terms it refers to open, external vocabularies, e.g.: license (Open Definition), funders (FundRef) and grants (OpenAIRE).

- **I3**. (Meta)data include qualified references to other (meta)data
    - Each referenced external piece of metadata is qualified by a resolvable URL.

2.4 Increase data re-use (through clarifying licenses)

- **R1**. (Meta)data are richly described with a plurality of accurate and relevant attributes
    - Each record contains a minimum of DataCite's mandatory terms, with optionally additional DataCite recommended terms and Zenodo's enrichments.
- **R1.1**. (Meta)data are released with a clear and accessible data usage license
    - License is one of the mandatory terms in Zenodo's metadata, and is referring to an Open Definition license, but within TeleRehaB DSS restricted access will be chosen for the patient data.
    - Data downloaded by the users is subject to the license specified in the metadata by the uploader.
- **R1.2**. (Meta)data are associated with detailed provenance
    - All data and metadata uploaded is traceable to a registered Zenodo user.
    - Metadata can optionally describe the original authors of the published work.
- **R1.3**. (Meta)data meet domain-relevant community standards
    - Zenodo is not a domain-specific repository, yet through compliance with DataCite's Metadata Schema, metadata meets one of the broadest cross-domain standards available.

2.4.1 Data re-use; open and restricted/controlled access

(a) TeleRehaB DSS will only provide data that has been de-identified.

(b) the patients that will participate in the TeleRehaB DSS study will be fully informed and will provide their consent that access to their de-identified data can be granted in the future for specific scientific purposes.

Open Access can be provided for the following HOLOBALANCE datasets:

• anonymized data from WP5 (demographics, questionnaires, interviews, focus groups)

• anonymized data from WP2 (raw data from the IMU and pressure sensors, video of the whole protocol for annotation)

Restricted access will be required for the data generated from the 460 study participants assigned in the TeleRehaB DSS group in WP5 which include clinical information (at baseline), raw data from the IMU and pressure sensors (for the 80 participants assigned in the HOLOBALANCE group), processed data providing performance, satisfaction, motivation etc. metrics, assessment scales: mini BESTTest,

FGA, MoCA, CANTAB, RAPA, WHODAS 2.0, ABC, FES-I,EQ-5D-5L, Environmental Mobility Scale.

Use and re-use of the pilot dataset will be subject to the license under which the data objects were deposited.

TeleRehaB DSS will also consider to deposit the content under an embargo status and provide and end date for the embargo in order to explore exploitation possibilities which affect the availability of data for third parties and usually is not finalized before the end of the project and the outcomes of the pilot study. Zenodo was chosen since it supports restricted access to the data until the end of the embargo period; at which time, the content will become publicly available automatically.

Zenodo also supports Restricted Access, i.e. the option to deposit restricted files with the ability to share access with others if certain requirements are met. These files will not be made publicly available and sharing will be made possible only by the approval of the TeleRehaB DSS data management board.

3. Allocation of resources

3.1 Costs

Equipment purchase costs and staff costs for preparing the data to be deposited will be covered by the project. Thus, costs are minor since they include only server maintenance costs or zero in case data are uploaded in Zenodo.

3.2 Resources for long term preservation (costs and potential value, who decides and how what data will be kept and for how long, etc.)

These issues are defined in this document. Final decisions will be made by December 2024 from the defined Board.

4. Data security

Each component of the system that generating or consuming data will be associated with a set of usage policies that ensure both security and privacy. Zenodo is a general-purpose open repository developed under the European OpenAIRE program and operated by CERN. One of the main reasons for Zenodo adoption is because of the security it provides:

- The data centres are located on CERN premises and all physical access is restricted to a limited number of staff with appropriate training and who have been granted access in line with their professional duties.
- The servers are managed according to the CERN Security Baseline for Servers, meaning e.g. remote access to our servers are restricted to Zenodo staff with appropriate training, and the operating system and installed applications are kept updated with latest security patches via the automatic configuration management system Puppet.
- The CERN Security Team runs both host and network-based intrusion detection systems and monitors the traffic flow, pattern and contents into and out of CERN networks in order to detect attacks. All access to zenodo.org happens over HTTPS, except for static documentation pages which are hosted on GitHub Pages.
- Zenodo stores user passwords using strong cryptographic password hashing algorithms (currently PBKDF2+SHA512). Users' access tokens to GitHub and ORCID are stored encrypted and can only be decrypted with the application's secret key.
- Zenodo also employs a suite of techniques to protect your session from being stolen by an attacker when a user is logged in and run vulnerability scans against the application.

Moreover, during the project:

- Information will be kept in locked filing cabinets and password protected computers, in a room with restricted access at the premises of the 4 partners participating in the pilot.
- Back up will be kept in an external hard disc locked in another room.
- Data transfers will be done through SFTPs.

6. Ethical aspects

The data will be collected from consenting patients that will be fully informed that their de-identified and anonymized data may be used for research also after the end of the project. Details about the protocol and the information sheets and consent forms are provided in the Ethics Deliverables.

Informed consent for data sharing and long-term preservation should be part of the patients' recruitment and all relevant issues should be clearly explained in the information sheets.

More details about Ethical aspects will be provided in **D7.1 and D7.2 (WP7).**

7. Conclusions

The first version of the TeleRehaB DSS Data Management Plan is included in this deliverable. The DMP adopts FAIR Principles. Good research data management is a priority for TeleRehaB DSS since it can lead to knowledge discovery and innovation, and to subsequent data and knowledge integration and reuse.

The DMP is intended to be a living document in which information will be made available on a finer level of granularity through updates as the implementation of the project progresses and when significant changes occur. As a minimum, the DMP will be updated in the context of the periodic evaluation/assessment of the project.


TeleRehaB

A.1 Appendix

A.1.1 Appendix

AGREEMENT
FOR THE SHARING OF ANONYMISED DATA
WITH TELEREHAB DSS CONSORTIUM
IN RELATION TO
THE RESEARCH PROJECT
GRANT AGREEMENT: [<<INSERT>>]

This Agreement is made on [<< *date* >>]

BETWEEN:

(1) **[<< UNIVERSITY COLLEGE LONDON**, a body corporate established by Royal Charter with company number RC000631 and whose principal place of business is Gower Street, London WC1E 6BT, England (**TELEREHAB DSS CONSORTIUM**)>>];

AND

(2) **[<< name >>]** whose principal place of business is at [<< *address details* >>] (**Data Provider**).

BACKGROUND

A. TELEREHAB DSS CONSORTIUM is the lead organisation for the Research Project.

B. The Data Provider is the custodian of the Study Data.

C. The Data Provider wishes to make the Study Data, in anonymised form only, available to TELEREHAB DSS CONSORTIUM for use in connection with the Research Project.

D. This Agreement sets out the terms and conditions upon which the Data Provider will make available to TELEREHAB DSS CONSORTIUM, and TELEREHAB DSS CONSORTIUM will use, the Study Data.

KEY DETAILS

Research Project	[<< *Insert details of the research project* >>]

Study Data	[<< *Insert details of the data to be made available by the Data Provider to TELEREHAB DSS CONSORTIUM* >>]

EXECUTION

Signed for and on behalf of [<<Partner>>] of TELEREHAB DSS CONSORTIUM	
Signature	[<<Insert>>]
Name of Authorised Signatory	[<<Insert>>]
Date	[<<Insert>>]

Signed for and on behalf of the Data Provider [<<Provider>>]	
Signature	[<<Insert>>]
Name of Authorised Signatory	[<<Insert>>]
Date	[<<Insert>>]

IT IS AGREED:

1. **Definitions and interpretation**

In this Agreement (including the Background):

1.1. **Definitions**

1.1.1. The following words and expressions have the following meanings:

Confidential information	Any information or materials (whether in writing, electronic form, oral or otherwise) concerning the affairs of one party that the other party obtains or receives as a result of the discussions leading up to or the entering into or the performance of this agreement that is confidential in nature or is marked or identified as confidential at the time of disclosure. The study data shall not constitute the confidential information of either party;
Data breach	The accidental or unlawful destruction, loss, alteration, unauthorised disclosure of, or access to, the study data held by TeleRehaB DSS consortium from time to time;
Parties	TeleRehaB DSS consortium and the data provider;
Personal data	Any information relating to an identified or identifiable living individual;
Research project	The research project described on the first page of this agreement;
Results	All results, information and materials created from analysis of the study data in the course of carrying out the research project;
Study data	The data, in anonymised (i.e., stripped of any personal identifiers) form only, described on the first page of this agreement;
Study protocol	The study protocol for the research project created and signed off by TeleRehaB DSS consortium in its role as the lead institution for the research project; and
Transparency laws	- The European Data Governance and Amending Regulation (Eu) 2018, 1724 (Data Governance Act) - The regulation (EU) 2018/1807 of the European parliament and of the council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union - The directive (EU) 2019/1024 of the European parliament and of the Council of

20 June 2019 on open data and the re-use of public sector information

1.2 Interpretation

1.2.1. words importing the singular shall include the plural and vice versa, words importing a gender shall include all genders and words importing persons shall include bodies corporate, unincorporated associations and partnerships;

1.2.2. references to Clauses are references to clauses of this Agreement;

1.2.3. Clause headings are included for convenience only and shall not affect the interpretation of this Agreement;

1.2.4. any reference to persons includes natural persons, firms, partnerships, companies, corporations, associations, organisations, governments, states, governmental or state agencies, foundations and trusts (in each case whether or not having separate legal personality and irrespective of the jurisdiction in or under the law of which it was incorporated or exists);

1.2.5. a reference to a statute or statutory provision is a reference to that statute or statutory provision and to all orders, regulations, instruments or other subordinate legislation made under the relevant statute;

1.2.6. any reference to a statute, statutory provision, subordinate legislation, code or guideline (legislation)
is a reference to such legislation as amended and in force from time to time and to any legislation which re-enacts or consolidates (with or without modification) any such legislation; and

1.2.7. any phrase introduced by the terms including, include, in particular, for example or any similar expression shall be construed as illustrative and shall not limit the sense of the words preceding those terms.

2. PURPOSE AND SCOPE OF THIS AGREEMENT

2.1. The purpose of this Agreement is to govern the sharing of the Study Data by the Data Provider with TELEREHAB DSS CONSORTIUM in connection with the Research Project.

2.2. This Agreement, together with the documents referred to in it, constitutes the entire agreement between the Parties in respect of the sharing and use of the Study Data and supersede and replace all previous negotiations, agreements, and commitments with respect thereto with effect on and from the date of this Agreement.

3. PROVISION OF THE STUDY DATA

3.1. The Data Provider shall make available the Study Data to TELEREHAB DSS CONSORTIUM for use in connection with the Research Project.

3.2. The Data Provider shall ensure that the Study Data it makes available to TELEREHAB DSS CONSORTIUM:

(a) is made available to TELEREHAB DSS CONSORTIUM in the format agreed between the
Parties;

(b) is fully anonymised so that it does not contain any Personal Data; and

(c) may lawfully be used and shared on the basis contemplated by this Agreement without breaching the rights of any third party.

3.3. Where it is agreed between the Parties that a copy of the Study Data will be transferred to TELEREHAB DSS CONSORTIUM, the Data Provider shall transfer the Study Data to TELEREHAB DSS CONSORTIUM via an encrypted secure file transfer mechanism.

3.4. The Data Provider shall ensure that it has obtained all approvals, consents, licences and permits which are necessary for it to be able to share the Study Data, and for the Study Data to be used, for the purposes contemplated by this Agreement.

4. DATA PROTECTION

4.1. The Parties acknowledge that only fully anonymised Study Data is to be made available to TELEREHAB DSS CONSORTIUM under this Agreement and that TELEREHAB DSS CONSORTIUM is not expecting to receive from the Data Provider Study Data that includes any Personal Data.

4.2. For these purposes, this means that, whilst the Data Provider itself may be able to link the Study Data it provides to TELEREHAB DSS CONSORTIUM back to specific, identifiable individuals, in TELEREHAB DSS CONSORTIUM's hands TELEREHAB DSS CONSORTIUM will not hold all of the information needed to link the Study Data back to specific, identifiable individuals.

4.3. TELEREHAB DSS CONSORTIUM shall not use the Study Data to identify or contact the individuals to whom the Study Data relates.

5. USE OF THE STUDY DATA

5.1. TELEREHAB DSS CONSORTIUM shall:

(a) use the Study Data only in connection with the Research Project;

(b) use the Study Data in compliance with all applicable laws and the Study Protocol;

(c) not use the Study Data for any commercial purpose; and

(d) not sell any of the Study Data.

5.2. TELEREHAB DSS CONSORTIUM may make available the Study Data to other institutions that are involved in the Research Project for use by those institutions in connection with the Research Project. TELEREHAB DSS CONSORTIUM shall only make available the Study Data to any such institution where it has in place with the institution a written agreement that places obligations on the institution regarding its use of the Study Data that are no less onerous than the obligations placed on TELEREHAB DSS CONSORTIUM regarding its use of the Study Data under this Agreement.

[*<< Note: include this Clause* ***Error! Reference source not found.*** *only where T ELEREHAB DSS CONSORTIUM will be sharing the Study Data with other institutions involved in the Research Project. If the Study Data is not being shared then you should delete this Clause* ***Error! Reference source not found.*** *>>*]

6. SECURITY OF THE STUDY DATA

6.1. TELEREHAB DSS CONSORTIUM shall implement appropriate technical and organisational measures to help ensure the security of the Study Data it holds from time to time and to try and prevent Data Breaches.

6.2. TELEREHAB DSS CONSORTIUM shall notify the Data Provider without undue delay after becoming aware of any Data Breach and shall work in co-operation with the Data Provider to try and mitigate the impact of any Data Breach.

7. RETENTION OF THE STUDY DATA

7.1. Save as otherwise set out in Clause 7.2, TELEREHAB DSS CONSORTIUM shall:

(a) retain the Study Data in line with the retention period set out in the Study Protocol; and

(b) delete or destroy the Study Data in line with the requirements set out in the Study Protocol promptly following the end of the retention period.

7.2. Where this Agreement is properly terminated by the Data Provider in accordance with Clause 12.3, TELEREHAB DSS CONSORTIUM shall delete or destroy the Study Data promptly following the date of termination of this Agreement.

8. RIGHTS RELATING TO THE STUDY DATA

8.1. As between TELEREHAB DSS CONSORTIUM and the Data Provider, the Data Provider shall remain the owner of the Study Data and nothing in this Agreement shall prevent or restrict the Data Provider from using the Study Data for its own purposes.

8.2. TELEREHAB DSS CONSORTIUM shall not acquire any rights or interest in the Study Data other than the right to use the Study Data in accordance with the provisions of this Agreement.

8.3. TELEREHAB DSS CONSORTIUM shall have the right to use the Results as part of any publication or presentation relating to the Research Project (such as the interim results of, or final outcomes from, the Research Project). The Parties' respective contributions to the Research Project will be recognised in any such publication or presentation in accordance with standard best practice for academic research study publications/presentations.

9. CONFIDENTIALITY

9.1. Each Party shall hold in confidence all Confidential Information obtained from the other Party. Neither Party shall disclose to any third party any Confidential Information in relation to the other Party save as expressly permitted by this Agreement or with the prior express written permission of the other Party.

9.2. The provisions of Clause 9.1 shall not apply to any information which:

(a) is or becomes public knowledge other than by breach of this Clause 9;

(b) is already in the possession of a Party without restriction in relation to disclosure before the date of its receipt from the other Party; or

(c) is received from a third party who lawfully acquired or developed it and who is under no obligation restricting its disclosure.

9.3. A Party may disclose Confidential Information in relation to the other Party:

(a) to those of its officers, employees, professional advisers, parent or subsidiary companies, or agents or sub-contractors as may be reasonably necessary for the purpose of fulfilling its obligations under this Agreement or, in the case of professional advisors, for use in their professional capacity, provided that before any such disclosure that Party shall make such officers, employees, professional advisers, parent or subsidiary companies, or agents or subcontractors aware of its obligations of confidentiality under this Agreement and shall at all times procure compliance by those persons with them; or

(b) where such disclosure is required by any law, court order or competent regulatory authority.

9.4. Without prejudice to the other rights of the disclosing Party, in the event of unauthorised disclosure or use of its Confidential Information occurring directly or indirectly through disclosure made to the receiving Party, the receiving Party shall (as soon as it becomes aware of the same) notify the disclosing Party of such unauthorised disclosure and use all reasonable endeavours to assist the disclosing Party in recovering and preventing the use of, dissemination, sale or other disposal of such Confidential Information.

9.5. Unless required to do so by applicable laws, neither Party shall make public the details of the terms or the operation or circumstances of termination of this Agreement without the other Party's prior written consent.

9.6. The Data Provider shall provide all necessary assistance and co-operation as reasonably requested by TELEREHAB DSS CONSORTIUM to enable TELEREHAB DSS CONSORTIUM to comply with its obligations under the Transparency Laws and, in particular, shall provide copies of any information requested by TELEREHAB DSS CONSORTIUM within seven days of TELEREHAB DSS CONSORTIUM's request.

10. INDEMNITY

10.1. The Data Provider shall indemnify TELEREHAB DSS CONSORTIUM from and against all costs (including the cost of bringing or defending any legal action), damages, losses and expenses suffered or incurred by TELEREHAB DSS CONSORTIUM arising out of or in connection with any breach by the Data Provider of its obligations under this Agreement.

11. LIABILITY

11.1. TELEREHAB DSS CONSORTIUM acknowledges and agrees that the use of the Study Data is for research purposes only and that such use is made available by the Data Provider free of charge and, save as otherwise expressly set out in this Agreement, on an "as is" basis. Save as otherwise expressly set out in this Agreement, the Data Provider makes no representations, and gives no warranties or undertakings, in respect of the Study Data. All warranties, conditions, terms, undertakings and obligations implied by statute, common law, custom, trade usage, course of dealing or otherwise are hereby excluded to the extent permitted by applicable law.

11.2. To the extent permitted by applicable law, neither Party shall be liable to the other Party for any indirect or consequential loss or damage arising out of or in connection with this Agreement.

11.3. To the extent permitted by applicable law, TELEREHAB DSS CONSORTIUM shall not be liable to the Data Provider for any use of the Study Data by any third party, except to

the extent that such liability arises out of TELEREHAB DSS CONSORTIUM knowingly making the Study Data available to a third party in breach of its obligations under this Agreement.

1.4. To the extent permitted by applicable law, each Party's total liability to the other Party under or in connection with this Agreement (whether such liability arises under any statute or in contract, tort (including negligence) or otherwise) shall be limited to £ [<< *insert appropriate overall limit of liability amount* >>].

12. TERM AND TERMINATION

12.1. This Agreement shall commence on the date of this Agreement and shall continue until terminated in accordance with the terms and conditions of this Agreement.

12.2. This Agreement shall automatically terminate on the date that the Research Project is completed or terminated prior to its completion.

12.3. Either Party may terminate this Agreement with immediate effect by giving written notice of such termination to the other Party if the other Party commits a material breach of any of the terms of this Agreement and either that breach is not capable of remedy or, if the breach is capable of remedy, the other Party fails to remedy that breach within 60 days of being notified of the breach.

13. EFFECTS OF TERMINATION

13.1. The termination of this Agreement shall not prejudice or affect any right of action or remedy which shall have accrued up to the date of termination.

13.2. Clauses 1, 2.2, 4.3, 6 to 11 (inclusive) and 13 to 15 (inclusive) shall survive termination of this Agreement and shall continue to apply as shall any other provision which by its nature is intended to survive termination.

14. NOTICES

14.1. Any notice required by this Agreement to be given by either Party to the other shall be in writing and shall be delivered by hand or sent by recorded delivery post or e-mail to the other Party at the address set out below or otherwise notified by the other Party in accordance with this Clause 14 from time to time.

TELEREHAB DSS CONSORTIUM	**Data Provider**
Attention of: [<< *insert* >>]	Attention of: [<< *insert* >>]
Address: [<< *insert* >>]	Address: [<< *insert* >>]
Email: [<< *insert* >>]	Email: [<< *insert* >>]

14.2. Any notice served under this Agreement shall be deemed to have been received (i) if delivered by hand, immediately upon delivery during the other Party's usual business hours; (ii) if sent by recorded delivery post, three days following delivery; or (iii) if sent by e-mail, when it is actually received by the recipient.

15. GENERAL

15.1. Neither Party shall assign, novate, sub-contract or otherwise dispose of any or all of its rights and obligations under this Agreement without the prior written consent of the other Party.

15.2. A person who is not a party to this Agreement shall have no rights under the Contracts (EU COUNCIL DIRECTIVE 93/13/EEC of 5 April 1993 on Unfair terms in consumer contracts) to enforce any term of this Agreement.

15.3. This Agreement shall not be construed as giving rise to the relationship of principal and agent (save as otherwise expressly provided herein) or partnership or joint venture.

15.4. If any provision of this Agreement or the application thereof to any Party or circumstance shall be declared void, illegal or unenforceable, the remainder of this Agreement shall be valid and enforceable to the extent permitted by applicable law. In such event, the Parties shall use their best efforts to replace the invalid or unenforceable provision by a provision that, to the extent permitted by applicable law, achieves the purposes intended under the invalid or unenforceable provision.

15.5. No delay or failure by a Party in exercising or enforcing any right or remedy under the terms and conditions of this Agreement will be deemed to be a waiver of any such right or remedy, nor will that failure operate to bar the exercise or enforcement of such right or remedy at any future time.

15.6. A change to this Agreement will only be effective if it is recorded in writing and signed by an authorised representative of each of the Parties.

15.7. This Agreement may be executed in counterparts, each of which shall be deemed to be an original and all of which together shall be deemed to be one and the same instrument.

15.8. This Agreement and any dispute or claim arising out of or in connection with it or its subject matter or formation (including, without limitation, non-contractual disputes or claims) shall be governed by and construed in accordance with the laws of England and Wales.

15.9. Each Party irrevocably agrees that the courts of [<< Insert Grand Agreement's Law Country Reference >>] shall have exclusive jurisdiction to settle any dispute or claim arising out of or in connection with this Agreement and its subject matter or formation (including non-contractual disputes or claims).

This page was intentionally left blank.

End of Agreement