BMJ Open

A TECHNICAL VALIDATION PROTOCOL FOR REAL-WORLD MONITORING OF GAIT

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-050785
Article Type:	Protocol
Date Submitted by the Author:	02-Mar-2021

Complete List of Authors:	Mazzà, Claudia; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Alcock, Lisa; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Aminian, Kamiar; Ecole Polytechnique Federale de Lausanne, Laboratory of Movement Analysis and Measurement Becker, Clemens; Robert Bosch Gesellschaft für Medizinische Forschung Bertuletti, Stefano; University of Sassari, Department of Biomedical Sciences Bonci, Tecla; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Brown, Philip; Newcastle Upon Tyne Hospitals NHS Foundation Trust Brozgol, Marina; Tel Aviv Sourasky Medical Center, Center for the Study of Movement, Cognition and Mobility, Neurological Institute Buckley, Ellen; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Carsin, Anne-Elie; Instituto de Salud Global Barcelona Caruso, Marco; Politecnico di Torino Dipartimento di Elettronica e Telecomunicazioni; Politecnico di Torino, PolitoBIOMed Lab – Biomedical Engineering Lab Caulfield, Brian; University College Dublin, Insight Centre for Data Analytics, O'Brien Science Centre; University College Dublin, UCD School of Public Health, Physiotherapy and Sports Science Cereatti, Andrea; Politecnico di Torino Dipartimento di Elettronica e Telecomunicazioni Chiari, Lorenzo; University of Bologna, Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi»; University of Bologna, Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV) Chynkiamis, Nikolaos; Northumbria University Newcastle, Department of Sport, Exercise and Rehabilitation Ciravegna, Fabio; The University of Sheffield, Department of Computer Science; The University of Sheffield, INSIGNEO Institute for in silico Medicine Del Din, Silvia; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Eskofier, Björn; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Artificial

	Intelligence in Biomedical Engineering Evers, Jordi; McRoberts BV Garcia Aymerich, Judith; Instituto de Salud Global Barcelona Gazit, Eran; Tel Aviv Sourasky Medical Center, Center for the Study of Movement, Cognition and Mobility, Neurological Institute Hansen, Clint; University Medical Center Schleswig-Holstein Campus Kiel, Department of Neurology Hausdorff, Jeffrey; Tel Aviv Sourasky Medical Center, Center for the Study of Movement, Cognition and Mobility, Neurological Institute; Tel Aviv University, Department of Physical Therapy, Sackler Faculty of Medicine & Sagol School of Neuroscience Helbostad, Jorunn; Norwegian University of Science and Technology, Department of Neuromedicine and Movement Science Hiden, Hugo; University of Newcastle upon Tyne, School of Computing Hume, Emily; Northumbria University Newcastle, Department of Sport, Exercise and Rehabilitation Paraschiv-Ionescu, Anisoara; Ecole Polytechnique Federale de Lausanne, Laboratory of Movement Analysis and Measurement Ireson, Neil; The University of Sheffield, Department of Computer Science; The University of Sheffield, INSIGNEO Institute for in silico Medicine Keogh, Alison; University College Dublin, Insight Centre for Data Analytics, O'Brien Science Centre; University College Dublin, UCD School of Public Health, Physiotherapy and Sports Science Kirk, Cameron; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Kluge, Felix; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering Koch, Sarah; Instituto de Salud Global Barcelona Küderle, Arne; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering Maetzler, Walter; University Medical Center Schleswig-Holstein Campus Kiel, Department of Neurology Micó-Amigo, M. Encarna; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Mueller, Arne; Novartis Pharma AG, Novartis Institutes of Biomedical Research Neatrour, Isabel; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Niessen, Martjin; McRoberts BV Palmerini, Luca; University of Bologna, Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi»; University of Bologna, Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV) Pluimgraaff, Lucas; McRoberts BV Reggi, Luca; University of Bologna, Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV) Salis, Francesca; University of Sassari, Department of Biomedical Sciences Schwickert, Lars; Robert Bosch Gesellschaft für Medizinische Forschung Scott, Kirsty; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Sharrack, Basil; Sheffield Teaching Hospitals NHS Foundation Trust, Department of Neuroscience and Sheffield NIHR Translational Neuroscience BRC Sillen, Henrik; AstraZeneca Sweden, Digital Health R&D Singleton, David; University College Dublin, Insight Centre for Data Analytics, O'Brien Science Centre; University College Dublin, UCD School

	of Public Health, Physiotherapy and Sports Science Soltani, Abolfazi; Ecole Polytechnique Federale de Lausanne, Laboratory of Movement Analysis and Measurement Taraldsen, Kristin; Norwegian University of Science and Technology, Department of Neuromedicine and Movement Science Ullrich, Martin; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Computer Science Van Gelder, Linda; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Vereijken, Beatrix; Norwegian University of Science and Technology, Department of Neuromedicine and Movement Science Vogiatzis, Ioannis ; Northumbria University Newcastle, Department of Sport, Exercise and Rehabilitation Warmerdam, Elke; University Medical Center Schleswig-Holstein Campus Kiel, Department of Neurology Yarnall, Alison; Newcastle University Faculty of Medical Sciences, Translational and Clinical Research Institute; Newcastle Upon Tyne Hospitals NHS Foundation Trust Rochester, Lynn; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences; Newcastle Upon Tyne Hospitals NHS Foundation Trust
Keywords:	Multiple sclerosis < NEUROLOGY, Parkinson-s disease < NEUROLOGY, Hip < ORTHOPAEDIC & TRAUMA SURGERY, Chronic airways disease < THORACIC MEDICINE, Heart failure < CARDIOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

A TECHNICAL VALIDATION PROTOCOL FOR REAL-WORLD MONITORING OF GAIT

Mazzà Claudia[1,2], Alcock Lisa[3], Aminian Kamiar[4], Becker Clemens[5], Bertuletti Stefano[6], Bonci Tecla[1,2], Brown Philip[7], Brozgol Marina[8], Buckley Ellen[1,2], Carsin Anne-Elie[9], Caruso Marco[10,11], Caulfield Brian[12,13], Cereatti Andrea[10], Chiari Lorenzo[14,15], Chynkiamis Nikolaos[16], Ciravegna Fabio[1,17], Del Din Silvia[3], Eskofier Björn[18], Evers Jordi[19], Garcia-Aymerich, Judith[9], Gazit Eran[8], Hansen Clint[20], Hausdorff Jeffrey[8,21], Helbostad Jorunn[22], Hiden Hugo[23], Hume Emily[16], Paraschiv-Ionescu Anisoara[4], Ireson Neil[1,17], Keogh Alison[12,13], Kirk Cameron[3], Kluge Felix[18], Koch Sarah[9], Küderle Arne[18], Maetzler Walter[20], Micó-Amigo M. Encarna[3], Mueller Arne[24], Neatrour Isabel[3], Niessen Martjin[19], Palmerini Luca[14,15], Pluimgraaff Lucas[19], Luca Reggi[15], Salis Francesca[6], Schwickert Lars[5], Scott Kirsty[1,2], Sharrack Basil[25], Sillen Henrik[26], Singleton David[12,13], Soltani Abolfazi[4], Taraldsen Kristin[22], Ullrich Martin[18], Van Gelder Linda[1,2], Vereijken Beatrix[22], Vogiatzis, Ioannis[16], Warmerdam Elke[20], Yarnall Alison[3,7], Rochester Lynn[3,7]

Corresponding author:
Prof. Claudia Mazzà
c.mazza@sheffield.ac.uk
Pam Liversidge Building - Room F14
Mappin Street, Sheffield, S1 3JD, UK
Phone: +44 (0)114 2226073

1. INSIGNEO Institute for in silico Medicine, The University of Sheffield, Sheffield, UK
2. Department of Mechanical Engineering, The University of Sheffield, Sheffield, UK
3. Translational and Clinical Research Institute, Faculty of Medical Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne, UK
4. Laboratory of Movement Analysis and Measurement, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland
5. Robert Bosch Gesellschaft für Medizinische Forschung, Stuttgart, Germany
6. Department of Biomedical Sciences, University of Sassari, Sassari, Italy
7. Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
8. Center for the Study of Movement, Cognition and Mobility, Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel
9. Instituto de Salud Global Barcelona, Barcelona, Spain
10. Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Turin, Italy
11. PolitoBIOMed Lab – Biomedical Engineering Lab, Politecnico di Torino, Turin, Italy
12. Insight Centre for Data Analytics, O'Brien Science Centre, University College Dublin, Dublin, Ireland
13. UCD School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland
14. Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi», University of Bologna, Bologna, Italy
15. Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV), University of Bologna, Bologna, Italy
16. Department of Sport, Exercise and Rehabilitation, Northumbria University Newcastle, Newcastle upon Tyne, UK
17. Department of Computer Science, The University of Sheffield, Sheffield, UK

18. Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
19. McRoberts BV, The Hague, Netherlands
20. Department of Neurology, University Medical Center Schleswig-Holstein Campus Kiel, Kiel, Germany
21. Department of Physical Therapy, Sackler Faculty of Medicine & Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel
22. Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway
23. School of Computing, University of Newcastle upon Tyne, Newcastle upon Tyne, UK
24. Novartis Institutes of Biomedical Research, Novartis Pharma AG, Basel, Switzerland
25. Department of Neuroscience and Sheffield NIHR Translational Neuroscience BRC, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
26. Digital Health R&D, AstraZeneca, Sweden

Keywords

Gait analysis; wearable sensors; activity monitoring; users evaluation; mobility outcomes

Wordcount: 5092

1 ABSTRACT

Introduction

Existing mobility endpoints based on functional performance, physical assessments and patient self-reporting are insensitive and resource-intensive, limiting their utility in clinical practice. Wearable devices including inertial measurement units (IMUs) can overcome these limitations by quantifying digital mobility outcomes (DMOs) both during supervised structured assessments and in real-world conditions. The validity of IMU-based methods in the real-world, however, is still limited in patient populations. Rigorous validation procedures should cover the metrological verification of a chosen device, the validation of the algorithms for the DMOs computation specifically for the population of interest and in daily life situations, and the users' perspective on the device.

Methods and analysis

This protocol was designed to establish the technical validity and patient acceptability of the approach used to quantify digital mobility in the real-world by Mobilise-D, an EU-funded IMI consortium aiming at fostering regulatory approval and clinical adoption of DMOs.

After defining the procedures for the metrological verification of an IMU based device, the experimental procedures for the validation of algorithms used to calculate the DMOs are presented. These procedures include the assessment of mobility both in the laboratory and in the real-world in 120 participants from five groups: healthy older adults; chronic obstructive pulmonary disease (COPD), Parkinson's disease (PD), multiple sclerosis (MS), proximal femoral fracture (PFF), and congestive heart failure (CHF). DMOs extracted from the monitoring device will be compared to those obtained from different validated reference systems, chosen according to the different contexts of observation. Questionnaires and interviews will evaluate the users' perspective on the deployed technology and on the relevance of the assessment of mobility.

Ethics and Dissemination

The study, registered at ISRCTN (12246987), has already been granted ethics approval by local committees of the involved centres. The data and algorithms used as part of this study will be made publicly available.

Strengths and limitations of this study

- A multi-disciplinary approach was implemented to define a protocol for the validation of tools for mobility monitoring, covering aspects related to devices, algorithms and users.
- A set of rigorous quality insurance procedures have been established to allow for the creation of a high-quality annotated dataset to foster development in the field of digital mobility monitoring.

- Mobility data will be collected in the laboratory and in the real-world for five different cohorts of slow-walkers. and subsequently will be made publicly available at the end of the study
- The multi-stage and multi-device experimental procedures required by this validation study can be extremely challenging for the both the participants and the assessors.
- For the laboratory acquisitions, the level of agreement between the gold standard and the inertial sensor devices might be affected by the limitations associated to a restricted capture volume.

2 INTRODUCTION

The ability to move is a key contributor to physical, mental and social well-being, which is in line with the World Health Organizations' (WHO's) definition of health. [1] However, the study of mobility has received relatively little attention, except for diseases characterised by specific mobility dysfunction. The increasing longevity of the world's population together with prolonged survival of many patients with long term conditions means that more people are suffering from loss of mobility, which in turn is a major determinant of loss of independence. [2,3] This has a considerable and growing personal, societal and economic impact. Efforts to mitigate this loss of mobility are an increasing priority and promising interventions are now under investigation. Existing mobility endpoints based on performance, patient self-reporting and one-off assessment are resource-intensive and lack sensitivity, [4] which limits therapeutic development and clinical management. A novel approach is needed that is low cost, simple, accurate and that can be used in the real world, including the home and the community. Poor gait, especially slow walking, is a key determinant of mobility loss. It is associated with greater mortality, morbidity, cognitive decline, dementia and fall risk. [2,3] Quantifying gait related mobility outcomes, including features such as such as step/stride duration and their variability, walking speed and asymmetry features is well established in supervised instrumented assessments. Wearable devices including inertial measurement units (IMUs) that allow digital mobility outcomes (DMOs) to be described, are leading the transition from laboratory-based assessment of mobility (mobility capacity), to continuous, unsupervised monitoring of mobility in daily life conditions (mobility performance). Nonetheless, the validity of IMU-based methods to characterise real-world mobility, and gait in particular, is still limited, especially in populations suffering from pathological conditions. This is because measuring real-world gait is far from simple or straightforward. In addition, complex factors arise from multiple sources that influence outcome measures, including disease characteristics, patient specific habits, environment/context and the purpose of walking. All these factors limit the validity of existing algorithms developed to quantify targeted DMOs. [5] Additionally, validation should include simultaneous evaluation [6] of the participants perception and acceptability of the device [7] as well as aspects related to wearability and usability. [7,8] Finally, a separate assessment of the metrological performance of the sensors contained in the adopted device is required. All these validation steps need

to be taken before the DMOs and the associated technologies can be effectively used for clinical [4] and regulatory [10] purposes (see Figure 1 for summary).

[Figure 1 around here]

In this paper, we present the comprehensive, multi-stage protocol deployed in the technical validation study (TVS) that we have developed as part of the IMI2-JU-funded Mobilise-D project (Number 820820 [11]), that aims to validate a new digital method for remote monitoring of mobility. In particular, this technical validation study aims to describe the multi-stage protocol to: a) verify the metrological performance of the sensors included in a IMU based monitoring device, using a procedure that could be replicated on any device; b) establish the validity and reliability [12] of the DMOs estimated by the algorithms using data from an IMU based device, taking into accounts the effects of populations (e.g., healthy adults, patients with various conditions), locomotor activities (simple straight walking versus complex walking tasks), contexts (lab based versus real-world), durations (device wearing time, DMOs hourly and daily fluctuations, etc.), and contextual confounding factors (such as location of walks, weather, use of walking aids, etc.); and, c) establish participants' and assessors' opinions on the usability and acceptability of the monitoring devices that will be deployed.

3 METHODS

The monitoring device that will be used in this study to collect mobility data is the DynaPort MM+ IMU (McRoberts, Table 1). In a trade-off between usability, accuracy and ability to provide DMOs for both at step and stride level, this is attached to the lower back via an elastic waistband and Velcro strap. The validity of algorithms to accurately estimate DMOs in the real-world will be investigated in 120 participants including: healthy older adults (HA) and in five clinical cohorts (chronic obstructive pulmonary disease, COPD; Parkinson's disease, PD; multiple sclerosis, MS; proximal femoral fracture, PFF; and congestive heart failure, CHF), chosen as presenting a variety of gait and mobility features. [4] The usability and acceptability of the device from the perspective of the participants and the assessors involved in the study will be established via interviews and questionnaires.

3.1 Verification of the metrological performances of the device

Verifying the performance of a device needs a robust and comprehensive metrological characterisation of all the sensors that it embeds. This requires a series of standardised procedures (spot-checks) to be implemented to ensure accuracy of raw data that will be used as input to the algorithms. Table 1 shows the sensing characteristics of the device used in this study.

Table 1 Characteristics of the sensors included in the DynaPort MM+ device

Sensor	Sampling Frequency	Sensor Range	Sensor resolution
Tri-axial accelerometer	100 Hz	± 8g	1 mg (at ± 8g)
Tri-axial gyroscope	100 Hz	± 2000 dps	70 mdps (at ± 2000 dps)
Barometer	1 Hz	300-1100 hPa	0.01 hPa
Temperature sensor	1 Hz	˚C	0.1 ˚C

According to Institute of Electrical and Electronics Engineers *Standard for Sensor Performance Parameter Definitions ((IEEE 2700-2017* [13]) a number of parameters are needed to characterise the metrological performance of the accelerometer, gyroscope, and magnetometer included in an IMU. These parameters can be computed under both static and dynamic conditions. The main noise parameters used are the first (mean value) and second (variance) order statistics, and the root Allan variance parameters of noise. [14,15] The parameters related to the first and second order statistics of noise can also be estimated by means of a short static acquisition (the minimum length of each acquisition is defined by IEEE standard for each sensor). These short static acquisitions can be simply performed using a plastic cube, where the device can be properly secured and then each of the three sensors' coordinate axes x, y and z are in turn aligned with the direction of gravity (g) as well as its opposite direction (six combinations, Figure 2). The root Allan variance parameters are instead computed over a long static acquisition. [16] Typically, the acquisitions are performed over a period of 4-8 hours. [14,15] All the static acquisitions should be carried out at a constant temperature of 25°C.

In comparison to a static acquisition, characterising the dynamic metrological performance of the sensors embedded in an IMU is less straightforward, since the metrological standards provided by IEEE describe a sequence of operations requiring an expensive and complex testing instrumentation. However, various alternatives have been proposed in the literature. The accuracy of a gyroscope can be quantified during a single-axis rotation by a known angle by computing the ideal angular velocity and comparing it to the average measured angular velocity. [17,18] This procedure should be performed using a rotation plate with a rotating speed comparable to what is encountered during human gait (~200 deg/s) (Figure 2).

[Figure 2 around here]

The tests described above will be performed on thirty-five different DynaPort devices deployed in the study. This will allow conformity with manufacturer indications to be verified, highlight the need for sensor recalibrations and provide benchmark standards for any device with equivalent sensing capacity. In turn, this will allow any device with an equivalent or superior solution to be used, in order to facilitate broader adoption of validated algorithms and cope with a continuously changing hardware landscape.

3.2 Protocol for Validation of the algorithms

Several algorithms to detect the DMOs from a single device have been implemented according to agreed definitions [19] and based on existing literature. At this time, these algorithms are being concurrently validated using the approach described by Bonci et al. [6] using pre-existing datasets, which mostly include lab-based observations only. Following this selection process, the best performing algorithms will be assessed using the data captured with the protocol here described.

3.2.1 Participants

This multi-centre study is sponsored and coordinated by The Newcastle upon Tyne Hospitals NHS Foundation Trust, UK. Participants will be recruited in five sites across Europe : Tel Aviv Sourasky Medical Center, Israel (ethics approval granted by the Helsinki Committee, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel, 0551-19TLV), Robert Bosch Foundation for Medical Research, Germany (ethics approval granted by the ethical committee of the medical faculty of The University of Tübingen, 647/2019BO2), University of Kiel, Germany (ethics approval granted by the ethical committee of the medical faculty of Kiel University, D438/18), The Newcastle upon Tyne Hospitals NHS Foundation Trust, UK and Sheffield Teaching Hospitals NHS Foundation Trust, UK (ethics approval granted by London – Bloomsbury Research Ethics committee, 19/LO/1507).

Table 2 Inclusion and exclusion criteria adopted for the different disease cohorts

Group	Inclusion criteria	Exclusion criteria
All groups	-able to walk 4 meters independently with or without walking aids -able to give informed consent -willingness to wear the sensor set-ups during the study -shoe size 36 (3 UK) or above -able to read and write in first language of the respective country -Montreal Cognitive Assessment (MoCA) >15 [20] -available for home /office visit during study period	-occurrence of any of the following 3 months prior to inclusion: myocardial infarction, hospitalization for unstable angina, stroke, coronary artery bypass graft (CABG), percutaneous coronary intervention (PCI), implantation of a cardiac resynchronization therapy device (CRTD) -current medical condition that could interfere with the patient's compliance
COPD	-≥45 years of age -Diagnosis of COPD (post-bronchodilator forced expiratory volume in the first second (FEV_1) to forced vital capacity (FVC) ratio <0.70) -clinical stability, defined as at least 4 weeks without antibiotics and/or oral corticosteroids to treat either a moderate or severe exacerbation -non-smokers, current or ex-smokers with a	-having undergone major lung surgery (e.g. lung volume reduction, lung transplant) -having a lung tumor -primary respiratory diseases other than COPD (e.g. asthma) - impaired mobility related to non-COPD causes, as judged by the investigator

	smoking history equivalent to at least 10 pack years (1 pack year = 20 cigarettes smoked per day for 1 year)	
PD	-aged 18+ years -Diagnosis of PD according to the Movement Disorders Society criteria [21]	-impaired mobility related to non-PD causes, as judged by the investigator
MS	-aged 18+ years -Diagnosis of MS based on the revised McDonald's criteria	-impaired mobility related to non-MS causes, as judged by the investigator
PFF	-65+ years of age -surgical treatment (fixation or arthroplasty) for a low-energy fracture of the proximal femur (ICD-10 diagnosis S72.0, S72.1, S72.2) as diagnosed on X-rays of the hip and pelvis within last 12 months	-impaired mobility related to non-PFF causes, as judged by the investigator
CHF	-≥45 years of age -Diagnosis of chronic heart failure NYHA class II-IV	- history of COPD ≥GOLD III - impaired mobility related to non-CHF causes, as judged by the investigator
HA	65+ years of age	

A convenience sample of 120 participants will be recruited via their clinical care team or research registers to represent the five disease cohorts (COPD, PD, MS, PFF, and CHF), as well as healthy older adults (HA). Twenty participants will be recruited for each cohort. Each cohort will be recruited across multiple sites to ensure generalizability (e.g. differing cultures and contexts). Inclusion and exclusion criteria, grouped by total cohort and disease cohort, are summarized in Table 2. The sample size of 120 has been defined according to Consensus-based Standards for the selection of health Measurement Instruments guidelines for measurement properties (COSMIN [22]). This sample size allows for 'excellent' methodological quality of non-inferiority studies, and is the one endorsed by the COSMIN checklist, a standardized tool for assessing the methodological quality of studies on measurement properties. [23]

All participants will give written informed consent prior to undergoing a clinic/laboratory-based session to record generic and disease-specific characterisations. This will include participant reported outcomes, assessments and medical notes review. The generic and cohort-specific clinical outcomes that will be collected are summarised in Table 3.

Table 3 Generic and cohort-specific clinical outcomes.

Cohort	**Generic Outcomes**
All	- Descriptive measures (age, sex, living arrangements, education) - Anthropometric measures (height, mass, shoe size, waist circumference) - Health status (comorbidities, number of falls and injuries in the 12 months prior to assessment, walking aid usage and current medication) - Montreal Cognitive Assessment to evaluate global cognition [20] - Visual Analogue Scale to measure pain during walking (0-10, from no pain to worse pain possible) - Function component of the Late-Life Function and Disability Instrument (LLFDI) to evaluate function and disability [24,25]
Cohort	**Cohort-specific clinical outcomes**
PD	- Movement Disorder Society (MDS)-sponsored revision of the Unified Parkinson's Disease (PD) Rating Scale (UPDRS), motor part [26]
MS	- Expanded Disability Status Scale (EDSS)* [27]
COPD	- 6-minute walk test - Recent spirometry test obtained from medical notes to characterise lung function* - COPD Assessment Test™ (CAT) [28]
PFF	- Short Physical Performance Battery (SPPB) – quiet standing balance task, a five times chair-raise test, and a 4m walk test at preferred gait speed
CHF	- 6-minute walk test - Kansas City Cardiomyopathy Questionnaire (KCCQ) * [29]

* denotes measures obtained from medical records if completed within 6 months prior to assessment

Data collection started across the above five sites in August 2020 and will continue till July 2021. All experimental procedures are constantly monitored and revised as needed to ensure full compliance with COVID19-related health and safety measures and for safeguarding of both the study participants and the assessors. The data collected as part of this protocol will be made publicly available together with the algorithms used to process the data.

3.2.2 Patient and Public Involvement

Patient and Public involvement and engagement (PPIE) has informed the design and conduct of this study. The protocol and patient facing documents were reviewed and changes implemented based upon reviewer feedback. We will work with our Patient Advisory Group (with members representing the patient cohorts involved in the study) to review study findings, data interpretation and, study reporting and dissemination, including co-design and presentation of dissemination materials for patients and the public. The learnings derived from these activities will further inform the work of the wider Mobilise-D project.

3.2.3 Experimental protocol

Performance of algorithms to determine walking-related DMOs is mostly affected by three factors: (i) the type of motor task (e.g., slow as opposed to fast, straight as opposed to curvilinear or inclined walking, etc.), (ii) the population of interest (e.g., healthy vs pathological gait), (iii) the context of observation (e.g., home vs outdoors). To accommodate these factors, we have developed a comprehensive, multi-stage protocol that includes a variety of tests conducted both in a laboratory context and in the real world (Table 4).

Table 4 Summary of the experimental protocol used for the validation of the algorithms in the laboratory and in the real world. INDIP= INertial module with DIstance Sensors and Pressure insoles.

Context of assessment	Reference Systems	Tested device	Mobility Tasks
Laboratory	Stereophotogrammetry	DynaPort MM+ INDIP	Structured mobility tasks and daily living activities
Real world (2.5 hours)	INDIP Mobile Phone with Aequora App Beacon	DynaPort MM+	Unsupervised real-world activities (including predetermined tasks)
Real world (7 days)	Mobile Phone with Aequora App Beacon	DynaPort MM+	Unsupervised daily living

Laboratory based assessment

Laboratory-based observations will be used to quantify validity and consistency within and between groups and different types of walking tasks under controlled ideal conditions. Structured and task-based mobility activities and a simulated daily activity session, mimicking habitual movements performed at home or at work will be included. The outcome of this comparison will provide the level of highest expected accuracy and minimum detectable changes for a given DMO.

Measurement tools

Reference system

A stereophotogrammetric (SP) system (100Hz) will be used as the gold standard in structured and simulated tests of daily activities to validate the DMOs calculated from the DynaPort raw data. A bespoke marker set will include four markers on each foot for detecting the gait events and four markers on the lower back device to track the displacement of the DynaPort device (Figure 3).

[Figure 3 around here]

To ensure quality and consistency in the SP data collection, accommodating different SP systems across sites, a spot check designed following the methodology proposed by Di Marco et al. [30] will be used, which will establish each system specific level of accuracy. A graphical user interface (GUI) for automated pre-processing of the SP data will ensure consistency in associated procedures (labelling, gap filling, etc.).

Tested devices

During this observation each participant will also be equipped with an additional multi-sensor system (INertial module with DIstance Sensors and Pressure insoles, INDIP) [31–33] and with the Dynaport device. The INDIP system (Figure 4) includes four inertial modules (one on the lower back, one on the non-dominant wrist and two on the feet), two distance sensors and two force-sensitive resistor pressure insoles including 16 force-resistive sensing elements (manufacturer 221e S.r.l., Italy). The INDIP has been designed to be used as a reference for real-world experiments, and in this phase of the protocol its performance will be validated against the SP system for the populations of interest. To this purpose, spatio-temporal parameters will be estimated exploiting the sensors redundancy and implementing previously validated sensor fusion algorithms. In particular, gait events will be detected using data from pressure insoles and inertial sensors separately and, then, combined for increasing robustness and detection accuracy (missed and extra events minimization). Secondly, spatial variables will be computed starting from feet inertial data using a Madgwick filter [34,35] combined with a zero-velocity update [36,37] and then calculating velocity and displacement using a direct and reverse approach. [38–41] Preliminary results from in-lab validation showed percentage errors of about 2% for gait speed as estimated during continuous walking, including both straight and curvilinear portions. [42]

[Figure 4 around here]

The lower back INDIP unit and the DynaPort will be rigidly attached to each other. The data from the SP (100Hz), INDIP (IMU and insoles, 100Hz, Distance sensor, 50Hz) and DynaPort (100Hz) systems will be synchronised using a hardware-based approach for the SP and the INDIP system, and timestamps to align recordings from the INDIP and the DynaPort.

Mobility tasks

Structured Mobility Tasks

Straight Walking: Straight walking is the most common test of walking. [43,44] The participant walks for a distance of 5m from a standing start and will be repeated at three different walking speeds: preferred, fast and slow (Figure 5 B).

Timed Up and Go (TUG): The TUG is a widely used clinical assessment of a person's mobility. [45] The participant is asked to sit in a chair, stand up, walk 3 m in a straight line, make a 180° turn, walk back to the chair, turn and sit down (Figure 5 A).

L Test: The participant is asked to sit in a chair, stand up, walk straight, turn 90° to the left around a cone, walk straight to the second cone, make a 180° turn to the left, walk straight before making a final 90° turn to the right and return to the chair to sit down (Figure 5 C). Besides being a clinically validated test, [46] the main purpose of including this test is the variation in curvilinear walking and the inclusion of different types of turns.

Surface Test: The participant walks around a defined circuit by turning around the cones (Figure 5 D). The circuit is completed twice, creating the longest walking bout out of all the tasks (approx. 20m).

Hallway Test: The participant walks along a 6m walkway stepping up and down a step positioned in the walkway. At the end of the walkway, the participant will complete a sharp 180° turn and walk back along the walkway (again stepping up and down off the step) until reaching the end point of the test (Figure 5 E).

Use of arm rests for the TUG and L-test and hand rails for the hallway test is permitted when needed.

[Figure 5 around here]

Daily living activities

These lab-based tasks will be used to simulate daily activities expected in the real life, similar to previous studies. [47] The participant starts by sitting in chair one and then executes a series of daily living tasks while sitting and moving around the room (see Figure 5F).

Real - world validation (2.5 hours observation)

This phase of the protocol will quantify validity and consistency across individuals and different types of walking tasks in the real world. It will be performed in a habitual environment (home/work/community) chosen by the participants. The duration of the observation has been established as trade-off between experimental, clinical and technical requirements.

Measurement tools

Reference systems

Participants will be asked to wear the INDIP, which in this phase of the protocol will be used as a reference system for the quantification of the DMOs provided by the single sensor algorithms, as applied to the DynaPort device data.

In order to quantify the effects of contextual confounding factors, the participants will also be provided with a system detecting outdoors walking, gradient of descent/ascent (walking uphill/downhill). The system is developed as a mobile Android application (Aeqora app) and the device selected was a Samsung S9 with Android 10. The app is composed of three parts: *(i)* the core tracker, *(ii)* the interface and *(iii)* the server infrastructure collecting data across users. The core tracker, adapted from a library

developed by the University of Sheffield, [48] utilises the mobile phone's internal sensors to compute the type of activity (e.g. walking) and intensity (e.g. cadence) to identify geo-located bouts of movement. It operates in the background and senses mobility features through a range of sensors (e.g. step counters, activity recognition, accelerometer, gyroscope, etc.) as well as from location services (GPS, network, Bluetooth, etc.). It collects the data and stores the raw sensor data into a local database in real time. A set of mechanisms have been developed to control access to this data, keep it secure, and regulate its use. First, no user identity information is sent within a single request as a token identifier is used. Additionally, a security layer is built based on Secure Socket Layer (SSL) and Transport Layer Security (TLS) 3.0 protocol to the data with scalable and efficient encryption algorithms. An SSL/TLS certificate is issued and used to establish identity and trust between server and client apps (desktop and mobile), ensuring privacy and security whenever communicating sensitive data.

Data collected during the experiments will be sent to a cluster of servers that uses algorithms to integrate the phone's data with contextual information about the locations where the participant will walk: where possible walks will be matched to OpenStreetMap [49] roads and paths, to remove GPS noise, the slope variation of each walk is computed on tiles, with a resolution of 5m within the UK (using Ordnance Survey Terrain 5) [50] and 30m in the other locations (using NASA's Shuttle Radar Topography Mission (SRTM) data [51], indoors and outdoors walking is recognised. Moreover, weather is associated with participant location based on the most proximate weather station.

The use of walking aids will also be monitored in this phase. For this purpose, a Bluetooth beacon (BlueBeacon Tag, BlueUp) will be attached to the walking aid and its activity will be detected by the phone's mobile tracker and saved by the app. The distance between the phone and the Beacon and data from the accelerometer contained in the Beacon will be integrated to determine when the aid is in use.

Mobility Tasks

To capture the largest possible range of activities during this assessment, participants will be guided by the following list of activities to be included: rise from a chair and walk to another room; walk to the kitchen and make a drink; walk up and down a set of stairs (if possible); walk outdoors (if possible, for a minimum of two minutes); if walking outside, walk up and down an inclined path. No supervision or structure to how these tasks should be completed will be given to the participants.

Real - world validation: seven days monitoring

This observation will quantify the effects of device wearing time, hourly and daily fluctuations of DMO's, and contextual confounding factors (such as location of the walk, weather, type of housing, etc.).

Measurement tools

The participants will be asked to wear the DynaPort, and to carry a mobile phone equipped with the Aequora App. Bluetooth beacons will also be used to track the use of walking aids.

Mobility Tasks

Participants will be monitored continuously for seven days, without any specific instruction being provided, except for that of wearing the provided measurement tools.

3.3 Assessment of participants' and assessors' experience

This part of the study will evaluate the participants' and assessors' experience of using the monitoring device. For the participant's assessment, wear-time of the device during the 7 days monitoring will be collected as a primary measure of compliance. Following the period of the seven-day, free-living data collection, participants will complete two questionnaires to assess the acceptability of the device. The first is a 12-item questionnaire [52] investigating usability on a 5-point ordinal scale. The questions are simple and focus on the impact of using a wearable device on participants' feelings, comfort and the ease of use of the device. The second questionnaire is the Comfort Rating Scale, [53] a 6-item measure investigating the comfort of a wearable device on a 21-point ordinal scale from '0 – low agreement' to '20 – high agreement'.

A subset of participants from all recruiting sites and cohorts will complete a semi-structured interview. For this qualitative part of the study, sampling will continue until saturation, i.e. until no additional learning is identified from the data. The interview will explore participants' opinions on the use of wearable devices and digital technology in healthcare, experiences of managing their condition, experiences of technology, and opinions on data privacy associated with the use of technology in healthcare. Additionally, participants will be asked about their experiences of using the device, including comfort perceived usefulness and ease of use, barriers and facilitators, and any other usability experiences that they may have encountered. All interviews will be audio-recorded, transcribed verbatim and, where required, translated to English.

To assess the professionals' experience, assessors from each of the clinical sites will be asked to assess the usability of the device after completion of the data collection. They will be provided with three questionnaires: 1) The System Usability Scale [54] a commonly used, validated 10-item questionnaire that asks users to rate a device on a 5-point Likert scale from '1 strongly disagree' to '5 – strongly agree'. Questions focus on the ease of use of the device, and the integration of various functions within it; 2) The IBM Computer System Usability Questionnaire [55] (to assess the DynaPort software), is composed of 19 items and asks respondents to consider their interaction with a computer system on a 7-point Likert scale from the perspective of data collection; 3) A bespoke questionnaire designed specifically for the TVS to assess the acceptability and effectiveness of the training methods, procedures, and any other materials provided within the study. The questionnaire will ask respondents to rate their experiences on a 7-point Likert scale to determine whether any changes to the procedures and materials are required, and whether training was effective in preparing researchers to implement the assessment protocol as planned.

In addition, assessors will complete a semi-structured interview with the aim of exploring their experiences of the data collection process. Assessors will be asked about the use of the device (e.g., ease of use, intuitiveness, data collection and download procedures, etc.), training and materials provided prior to the commencement of the study, and barriers and facilitators to using the device. The topic guide and open-ended questions allow for new areas of conversation to emerge. All interviews will be audio-recorded and transcribed verbatim.

3.4 Data management

All data will be uploaded to a central platform "e-Science Central" [56] which provides data processing and storage functionality in accordance with principles of reproducible research. The underlying infrastructure complies with ISO 27000 standards for Information Security Management Systems and is hosted on Amazon Web Service secure services cloud platform. Data will be integrated on the platform by means of implementation of a standardised file nomenclature system. At point of capture, each file will be labelled in standardised format. For source data, we will adhere to principles defined by the US Food and Drug Administration [57] for making them attributable, legible, contemporaneous, original, and accurate (ALCOA+).

3.5 Data analysis plan

3.5.1 Verification of the device

Mean and standard deviation readings of the accelerometer, gyroscope, and magnetometer signals captured during static acquisition will be used to assess the reliability of the manufacturer sensor calibration and to detect the presence of abnormal spikes in the sensor signals. Data from long static acquisitions will be used to confirm the stability of sampling frequency, the duration of the battery, and to estimate the Allan deviation (bias instability) of the gyroscope readings over time. Errors of gyroscope readings will be assessed using mean and standard deviation of nominal, measured and relevant errors for angular velocity values during the dynamic acquisitions.

3.5.2 Validation of the algorithms

The data analysis will determine criterion validity (including selected criterion (concurrent) validity and reliability metrics) of the primary and secondary DMOs listed in Table 5. Table 6 summarises the statistical tools that will be used to quantify each of the DMOs. All statistical analyses will be performed using the statistical analysis toolbox of Matlab R2018a. Continuous variables (e.g. cadence, real-walking-speed) will be summarized with descriptive statistics for the values obtained within walking bouts (mean and standard deviation). In addition, the mean, minimum, maximum, standard deviation, median, interquartile range (IQR) and Root Mean Square Error (RMSE) of DMOs over all available walking bouts will be presented. Categorical variables (e.g. laterality of initial contacts) will be

summarized with frequency counts and percentages. Confidence Interval 95% (CI) will be provided for the Interclass Correlation Coefficients (ICC).

For specific DMOs (e.g. events, etc.), True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN) will be identified and used to calculate the following performance metrics:

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

$$\text{Positive Predicted Value} = \frac{\text{TP}}{\text{TP} + \text{FP}},$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}},$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \text{ and}$$

$$\text{F1} - \text{score} = 2 * \frac{\text{Positive Predicted Value} * \text{Sensitivity}}{\text{Positive Predicted Value} + * \text{Sensitivity}}.$$

Criterion validity will be characterised by evaluating the absolute and relative differences (error) between the DMOs quantified with the single device and those derived from the reference systems. Then, the mean, standard deviation and maximum of all absolute errors will be reported for each walking bout. Limits of agreement between single sensor and reference system DMOs will be quantified. In addition, statistically significant differences between single sensor and reference system DMOs will be evaluated via parametric (t-test) or non-parametric (Wilcoxon signed-rank test) tests, depending on the distribution of the data.

Concurrent validity between the DMOs quantified by the single device and those derived from the reference systems will be evaluated by quantifying the Intraclass Correlation Coefficient (ICC(2,1)).

All results will be presented separately by cohort (e.g. PD) and subgroup (i.e., subgroups of the cohorts) stratified by average stride gait speed (e.g. fast speed: walking speed > 1 m/s, medium speed: walking speed between 0.5m/s and 1 m/s, slow speed: walking speed < 0.5 m/s) [2].

Table 5 List of primary and secondary DMOs that will be analysed as part of the TVS.

Variables	DMOs (units)	Definition
Walking Bout (WB) *A walking sequence containing at least two consecutive strides of both feet. Start and end of a WB are determined by a resting period or any other activity (non-walking period).*	Number of WBs (count)	Based on the identification of gait as an activity (yes/no) to a sample level of 0.1 s
	WB Start (s)	Start of WB
	WB End (s)	End of WB
	WB Duration (s)	Time between start and the end of WB
Stride/ Step Duration (SD) *Refers to the duration (time intervals) of strides, calculated as the time in between two non-consecutive (alternate) initial contacts.*	Stride Duration (s)	Duration between two non-consecutive (alternate) initial contact events
	Step Duration (s)	Duration between two consecutive initial contact events

Cadence (CE)	Cadence (steps/minute)	Steps performed within a minute
Stride Length (SL)	Mean Stride Length (m)	Average stride length within a WB
Real-World - Walking-Speed (RWS)	Gait speed (m/s)	Velocity, ratio between displacement covered within a WB and time to cover it
Turning	Number of Turns	Based on the identification of turns (yes/no) to a sample level of 0.1 s
	Turn Start (s)	Start of each turn within the WB
	Turn End (s)	End of each turn within the WB
	Turn Duration (s)	Time between the start and the end of the turns within the WB
	Maximal Turn Angle (deg)	Maximal angle achieved in the turn
Height Estimation	Elevation Change (m)	Difference between the minimal and maximal height or elevation for the complete walking bout detected
Left/Right Identification	Laterality (label)	Left or Right category, indicating the foot with which the initial contact is performed
Secondary Outcomes (SO)	Number of Final Contact Events (counts)	Correct identification of Final Contact events
	Final Contact Event (s)	Instant of time at which each final contact event is performed within a walking bout
	Swing Phase Duration (s)	Time between the last contact of the current footfall and the first contact of the next footfall on the same foot
	Stance Phase Duration (s)	Time in between the first contact and the last contact of two consecutive footfalls on the same foot
	Variability of: Step Time, Stride Time, Swing Time, Stance Time Stride Velocity Stride Length (same units as variable)	St. Dev. and Coefficient of Variation of step time, of stride time, of swing time, of stance time, of stride velocity and stride length within a WB
	Asymmetry of: Step Time, Stride Time, Swing Time, Stance Time (same units as variable)	Asymmetry evaluated as difference between right and left steps or strides for step time, of stride time, of swing time and of stance time within a WB

Table 6. List of statistical analyses and performance metrics that will be used for the various DMOs. Performance metrics and criterion validity are those that will be used to compare DMOs obtained from a single device versus those obtained from the reference system. The types of plots listed in the table will be used to visualise performance metrics and to support interpretation of the results.

	PERFORMANCE METRIC					CRITERION VALIDITY							PLOTS		
DMO	**Sensitivity**	**Positive Predictive Value**	**Accuracy**	**Specificity**	**F1-score**	**Error (absolute & relative)**	**SD & Max of Error**	**Root Mean Squared Error**	**Precision**	**Concurrent validity**	**Significant Difference**	**Limit of Agreement**	**Bland Altman Plots**	**Scatter Plots for Correlation**	**Histogram Plots**
Number of Walking Bouts	✓	✓	✓	✓	✓										
WB Start						✓	✓	✓	✓				✓	✓	✓
WB End						✓	✓	✓	✓				✓	✓	✓
WB Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Initial Contact Events	✓	✓			✓	✓									
Step Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Stride Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Final Contact Events	✓	✓			✓	✓									
Mean Stride Length						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cadence						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Elevation Change						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Walking Speed						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Number of Turns	✓	✓	✓	✓	✓										
Turn Start						✓	✓	✓	✓				✓	✓	✓
Turn End						✓	✓	✓	✓				✓	✓	✓
Turn Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Maximal Turn Angle						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Laterality						✓	✓	✓	✓	✓					
Sequence Order						✓	✓	✓							

3.5.3 Participants' and assessors' experience

Participant and professional questionnaire data will be analysed using descriptive statistics. Interviews will be analysed using thematic analysis. [58] Transcripts will be deductively examined for the presence of themes related to the acceptability of the monitoring device to participants (i.e. comfort, interference with daily living) based on previous literature. Specifically, perceived usefulness, comfort, and ease of use are critical factors of usability thus, these will be the categories examined within the transcripts. Regarding participants use of technology to manage their healthcare condition an inductive approach will be taken. A list of codes relevant to the question will be generated and then refined by grouping them into potential themes.

4 CONCLUSIONS

The technical validation protocol provides the stakeholders, the first comprehensive, multi-modal solution to validate real-world mobility. The solution is based on a single device and covers all complementary aspects related to the device, algorithms to derive DMOs and the user perspective. Notably, the protocol described here has been accepted by the European Medical Agency,[59] as part of the Mobilise-D process for the regulatory qualification of real-world mobility performance biomarkers in Parkinson's disease. [10]

Authors' contributions:

Manuscript initial drafting: CM, TB, SDD, AC, LA, AK, KS, SB, FC. Design and deployment of experimental tools and protocols: KS, LS, CM, TB, EB, FC, NI, BS, LV, IV, EH, NC, FS. AC, MC, SB, LRO, AY, CK, EA, LS, CB, WM, CH, EW, JHA, EG, MB; J, MN, KT, BV, JHE, AK, DS, AM. Data analysis and algorithm development: SDD, EA, HS, AS, BE, FK, AK, MU, AC, JG, SK, LC, LP, LR, HH, EG. Draft revisions: LRO, KA, BE, JHA, BV, HS, AM, JHO, PB, FK, BC, AK, LC, LP, EB, KS, SK, LP, HH, WM, NC, BS, AI, IN Study design and coordination: CM, LRO, AM, SDD, AC. All authors have read and approved the final manuscript.

Competing interests statement

Fabio Ciravegna is CEO and shareholder of Aeqora. Henrik Sillén is an employee of AstraZeneca. Bjoern M Eskofier is co-founder and owns shares of Portabiles HealthCare Technologies GmbH. McRoberts is the manufacturer of the DynaPort. Martijn Niessen, Jordi Evers, and Lucas Pluimgraaff are employees of McRoberts. Arne Mueller is an employee of Novartis and holds stock in Novartis. Lars Schwickert and Clemens Becker are consultants of Philipps Healthcare, Bosch Healthcare, Eli Lilly, Gait-up. Jeffrey M. Hausdorff reports having submitted a patent for assessment of mobility using wearable sensors in 400 Parkinson's disease; the intellectual property rights 401 are held by the Tel Aviv Medical Center. Luca Palmerini and Lorenzo Chiari are co-founders and own shares of mHealth Technologies.

Funding

This study was co-funded by the European Union's Horizon 2020 research and innovation programme and EFPIA via the Innovative Medicine Initiative 2 (Mobilise-D project, grant number IMI22017-13-7-820820). The views expressed are those of the authors and not necessarily those of the IMI, the European Union, the EFPIA, or any Associated Partners. We acknowledge the support of Grünenthal GmbH via the funding of a PhD scholarship directly dedicated to the technical validation protocol.

We acknowledge support from the Spanish Ministry of Science and Innovation through the "Centro de Excelencia Severo Ochoa 2019-2023" Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program. SDD, AJY and LR are also supported by the Newcastle Biomedical Research Centre (BRC) based at Newcastle upon Tyne and Newcastle University. CM, BS and EB are also supported by the Sheffield Biomedical Research Centre (BRC) based at the Sheffield Teaching Hospital and the University of Sheffield. The work was also supported

by the NIHR/Wellcome Trust Clinical Research Facility (CRF) infrastructure at Newcastle upon Tyne Hospitals NHS Foundation Trust and the CRF at the Sheffield Teaching Hospital. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care or the funders.

References

1 Team WHOD of MH and SA and WHO and WHOD of MH and SAMH and WHOMHE an. *Mental health atlas 2005*. World Health Organization 2005.

2 Studenski S, Perera S, Patel K, *et al.* Gait Speed and Survival in Older Adults. *JAMA* 2011;**305**:50–8. doi:10.1001/jama.2010.1923

3 Perera S, Patel K V, Rosano C, *et al.* Gait Speed Predicts Incident Disability: A Pooled Analysis. *Journals Gerontol Ser A* 2016;**71**:63–71. doi:10.1093/gerona/glv126

4 Rochester L, Mazzà C, Mueller A, *et al.* A Roadmap to Inform Development, Validation and Approval of Digital Mobility Outcomes: The Mobilise-D Approach. *Digit Biomarkers* 2020;**4**:13–27. doi:10.1159/000512513

5 Del Din S, Godfrey A, Mazzà C, *et al.* Free-living monitoring of Parkinson's disease: Lessons from the field. *Mov Disord* 2016;**31**:1293–313. doi:10.1002/mds.26718

6 Bonci T, Keogh A, Din S Del, *et al.* An objective methodology for the selection of a device for continuous mobility assessment. *Sensors (Switzerland)* 2020;**20**:1–16. doi:10.3390/s20226509

7 Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. *Ind Manag Data Syst* 2015;**115**:1704–23. doi:10.1108/IMDS-03-2015-0087

8 Jia Y, Wang W, Wen D, *et al.* Perceived user preferences and usability evaluation of mainstream wearable devices for health monitoring. *PeerJ* 2018;**2018**:1–17. doi:10.7717/peerj.5350

9 Puri A, Kim B, Nguyen O, *et al.* User acceptance of wrist-worn activity trackers among community-dwelling older adults: Mixed method study. *JMIR mHealth uHealth* 2017;**5**. doi:10.2196/mhealth.8211

10 Viceconti M, Penna SH, Dartee W, *et al.* Toward a regulatory qualification of real-world mobility performance biomarkers in parkinson's patients using digital mobility outcomes. *Sensors (Switzerland)* 2020;**20**:1–13. doi:10.3390/s20205920

11 Mobilise-D website. https://www.mobilise-d.eu/ (accessed 18 Feb 2021).

12 Bartlett JW, Frost C. Reliability, repeatability and reproducibility: Analysis of measurement errors in continuous variables. *Ultrasound Obstet Gynecol* 2008;**31**:466–75. doi:10.1002/uog.5256

13 Systems M, Development S, Devices E. *IEEE Standard for Sensor Performance Parameter Definitions*. 2014.

14 El-Sheimy N, Hou H, Niu X. Analysis and modeling of inertial sensors using allan variance.

IEEE Trans Instrum Meas 2008;**57**:140–9. doi:10.1109/TIM.2007.908635

15 Hussen AA, Jleta IN. Low-Cost Inertial Sensors Modeling Using Allan Variance. *Int Sch Sci Res Innov* 2015;**9**:1069–74.

16 IEEE Aerospace and Electronic Systems Society (institution). *IEEE Standard Specification Format Guide and Test Procedure for Single-Axis Laser Gyros*. 1995.

17 Ferraris F, Gorini I, Grimaldi U, *et al.* Calibration of three-axial rate gyros without angular velocity standards. *"Sensors Actuators, A Phys* 1994;**42**:446–9. doi:10.1016/0924-4247(94)80031-6

18 Stančin S, Tomažič S. Time- and computation-efficient calibration of MEMS 3D accelerometers and gyroscopes. *Sensors (Switzerland)* 2014;**14**:14885–915. doi:10.3390/s140814885

19 Kluge F, Din S Del, Cereatti A, *et al.* Consensus based framework for digital mobility monitoring. Published Online First: 2020. doi:https://doi.org/10.1101/2020.12.18.20248404

20 Nasreddine ZS, Phillips NA, Bédirian V, *et al.* The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc* 2005;**53**:695–9. doi:https://doi.org/10.1111/j.1532-5415.2005.53221.x

21 Postuma RB, Berg D, Stern M, *et al.* MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2015;**30**:1591–601. doi:10.1002/mds.26424

22 COSMIN. https://www.cosmin.nl/ (accessed 18 Feb 2021).

23 COSMIN Study Design checklist for Patient-reported outcome measurement instruments. https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf (accessed 18 Feb 2021).

24 Jette AM, Haley SM, Coster WJ, *et al.* Late Life Function and Disability Instrument: I. Development and Evaluation of the Disability Component. *Journals Gerontol Ser A* 2002;**57**:M209–16. doi:10.1093/gerona/57.4.M209

25 Haley SM, Jette AM, Coster WJ, *et al.* Late Life Function and Disability Instrument: II. Development and Evaluation of the Function Component. *Journals Gerontol Ser A* 2002;**57**:M217–22. doi:10.1093/gerona/57.4.M217

26 Goetz CG, Tilley BC, Shaftman SR, *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord* 2008;**23**:2129–70. doi:https://doi.org/10.1002/mds.22340

27 Kurtzke JF. Rating neurologic impairment in multiple sclerosis. *Neurology* 1983;**33**:1444 LP –

1444. doi:10.1212/WNL.33.11.1444

28 Gupta N, Pinto LM, Morogan A, *et al.* The COPD assessment test: a systematic review. *Eur Respir J* 2014;**44**:873 LP – 884. doi:10.1183/09031936.00025214

29 Green CP, Porter CB, Bresnahan DR, *et al.* Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. *J Am Coll Cardiol* 2000;**35**:1245–55. doi:10.1016/S0735-1097(00)00531-3

30 Di Marco R, Rossi S, Castelli E, *et al.* Effects of the calibration procedure on the metrological performances of stereophotogrammetric systems for human movement analysis. *Meas J Int Meas Confed* 2017;**101**:265–71. doi:10.1016/j.measurement.2016.01.008

31 Bertuletti S, Cereatti A, Comotti D, *et al.* Static and dynamic accuracy of an innovative miniaturized wearable platform for short range distance measurements for human movement applications. *Sensors (Switzerland)* 2017;**17**:1–15. doi:10.3390/s17071492

32 Salis F, Bertuletti S, Caruso M, *et al.* Multi-sensor integration and data fusion for enhancing gait assessment in and out of the laboratory. *Gait Posture* 2019;**74**:34. doi:10.1016/j.gaitpost.2019.07.493

33 Bertuletti S, Della Croce U, Cereatti A. A wearable solution for accurate step detection based on the direct measurement of the inter-foot distance. *J Biomech* 2019;**84**:274–7. doi:https://doi.org/10.1016/j.jbiomech.2018.12.039

34 Caruso M, Sabatini AM, Knaflitz M, *et al.* Orientation Estimation Through Magneto-Inertial Sensor Fusion: A Heuristic Approach for Suboptimal Parameters Tuning. *IEEE Sens J* 2020;:1. doi:10.1109/JSEN.2020.3024806

35 Caruso M, Sabatini AM, Knaflitz M, *et al.* Accuracy of the Orientation Estimate Obtained Using Four Sensor Fusion Filters Applied to Recordings of Magneto-Inertial Sensors Moving at Three Rotation Rates. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019. 2053–8. doi:10.1109/EMBC.2019.8857655

36 Skog I, Nilsson J, Händel P. Evaluation of zero-velocity detectors for foot-mounted inertial navigation systems. In: *2010 International Conference on Indoor Positioning and Indoor Navigation*. 2010. 1–6. doi:10.1109/IPIN.2010.5646936

37 Peruzzi A, Della Croce U, Cereatti A. Estimation of stride length in level walking using an inertial measurement unit attached to the foot: A validation of the zero velocity assumption during stance. *J Biomech* 2011;**44**:1991–4. doi:10.1016/j.jbiomech.2011.04.035

38 Trojaniello D, Cereatti A, Pelosin E, *et al.* Estimation of step-by-step spatio-temporal

parameters of normal and impaired gait using shank-mounted magneto-inertial sensors : application to elderly , hemiparetic , parkinsonian and choreic gait. 2014;**44**:152.

39 Köse A, Cereatti A, Della Croce U. Bilateral step length estimation using a single inertial measurement unit attached to the pelvis. *J Neuroeng Rehabil* 2012;**9**. doi:10.1186/1743-0003-9-9

40 Trojaniello D, Cereatti A, Ravaschio A, *et al.* Assessment of gait direction changes during straight-ahead walking in healthy elderly and Huntington Disease patients using a shank worn MIMU. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2014. 2508–11. doi:10.1109/EMBC.2014.6944132

41 Trojaniello D, Cereatti A, Croce U Della. Gait direction of progression changes using shank worn MIMUs. In: *GBN Conference, Pavia*. 2014.

42 Salis F, Bertuletti S, Caruso M, *et al.* A novel multi-sensor system for gait assessment in real-world conditions: preliminary results. In: *Proceedings of GNB 2020*. 2020.

43 Zijlstra W, Hof AL. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture* 2003;**18**:1–10. doi:https://doi.org/10.1016/S0966-6362(02)00190-X

44 Ngueleu AM, Blanchette AK, Bouyer L, *et al.* Design and accuracy of an instrumented insole using pressure sensors for step count. *Sensors (Switzerland)* 2019;**19**:12–4. doi:10.3390/s19050984

45 Nightingale CJ, Mitchell SN, Butterfield SA. Validation of the Timed Up and Go Test for Assessing Balance Variables in Adults Aged 65 and Older. *J Aging Phys Act*;**27**:230–3. doi:10.1123/japa.2018-0049 10.1123/japa.2018-0049

46 Deathe AB, Miller WC. The L Test of Functional Mobility: Measurement Properties of a Modified Version of the Timed “Up & Go” Test Designed for People With Lower-Limb Amputations. *Phys Ther* 2005;**85**:626–35. doi:10.1093/ptj/85.7.626

47 Pham MH, Elshehabi M, Haertner L, *et al.* Validation of a Step Detection Algorithm during Straight Walking and Turning in Patients with Parkinson’s Disease and Older Adults Using an Inertial Measurement Unit at the Lower Back . Front. Neurol. . 2017;**8**:457.https://www.frontiersin.org/article/10.3389/fneur.2017.00457

48 Ciravegna F, Gao J, Ireson N, *et al.* Active 10: Brisk Walking to Support Regular Physical Activity. In: *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. New York, NY, USA: : Association for Computing Machinery 2019. 11–20. doi:10.1145/3329189.3329208

49 Open Street Map. https://www.openstreetmap.org (accessed 18 Feb 2021).

50 Ordnance Survey Terrain 5. https://www.ordnancesurvey.co.uk/business-government/products/terrain-5 (accessed 18 Feb 2021).

51 U.S. Releases Enhanced Shuttle Land Elevation Data. https://www2.jpl.nasa.gov/srtm/ (accessed 18 Feb 2021).

52 Rabinovich RA, Louvaris Z, Raste Y, *et al.* Validity of physical activity monitors during daily life in patients with COPD. *Eur Respir J* 2013;**42**:1205–15. doi:10.1183/09031936.00134312

53 Knight JF, Baber C, Schwirtz A, *et al.* The comfort assessment of wearable computers. *Proc - Int Symp Wearable Comput ISWC* 2002;**2002-Janua**:65–72. doi:10.1109/ISWC.2002.1167220

54 Brooke J. SUS: a “quick and dirty’usability". *Usability Eval Ind* 1196;**Jun 11**:189.

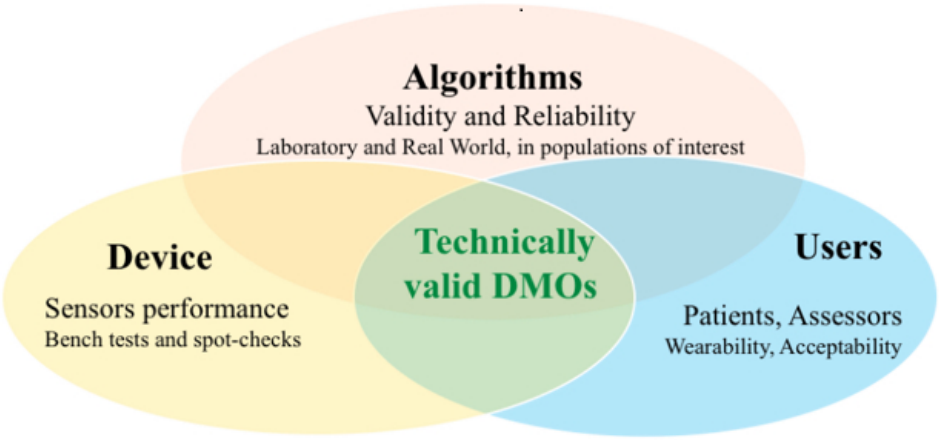
55 Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *Int J Human–Computer Interact* 1995;**7**:57–78. doi:10.1080/10447319509526110

56 Hiden H, Woodman S, Watson P, *et al.* Developing cloud applications using the e-Science Central platform. *Philos Trans R Soc A Math Phys Eng Sci* 2013;**371**. doi:10.1098/rsta.2012.0085

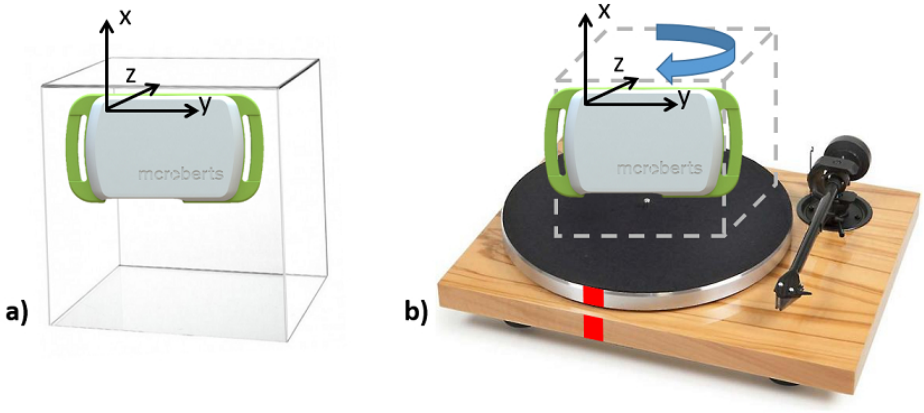
57 Data Integrity and Compliance With CGMP. Guidance for Industry. https://www.fda.gov/files/drugs/published/Data-Integrity-and-Compliance-With-Current-Good-Manufacturing-Practice-Guidance-for-Industry.pdf (accessed 18 Feb 2021).

58 Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101. doi:10.1191/1478088706qp063oa

59 Letter of support for Mobilise-D digital mobility outcomes as monitoring biomarkers. https://www.ema.europa.eu/en/documents/other/letter-support-mobilise-d-digital-mobility-outcomes-monitoring-biomarkers_en.pdf (accessed 18 Feb 2021).



Concurrent domains to be assessed as part of a technical validation of digital mobility outcomes (DMOs) obtained from a wearable device.



Testing configurations used during (a) short static (plexiglass cube and device) and (b) dynamic (turntable and device) acquisitions. Marks (in red) are applied on both on the turntable and on the base to identify start/end, which have to align for the dynamic tests.

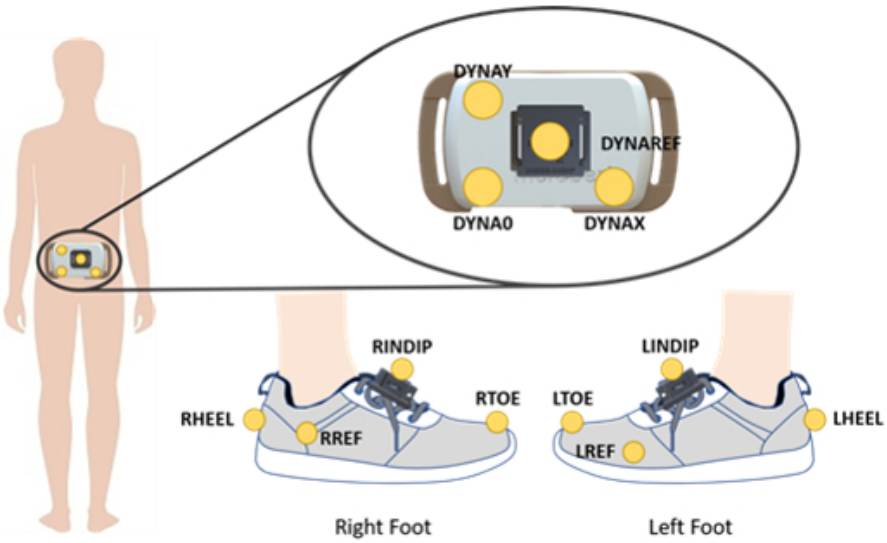
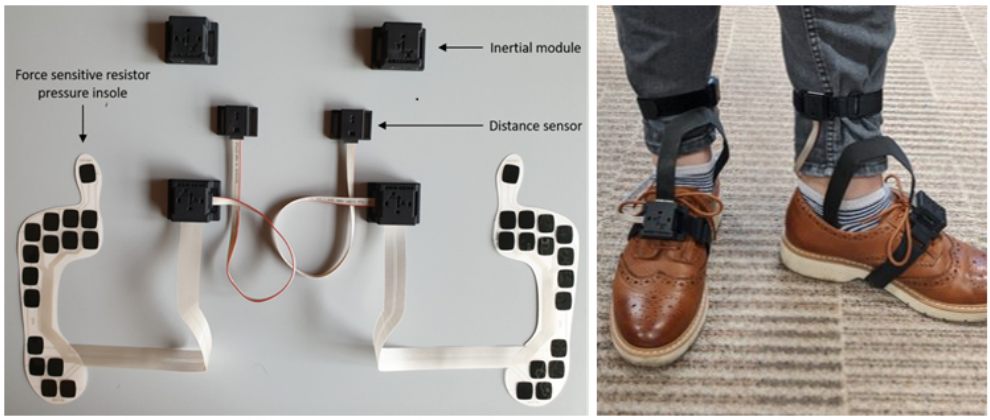
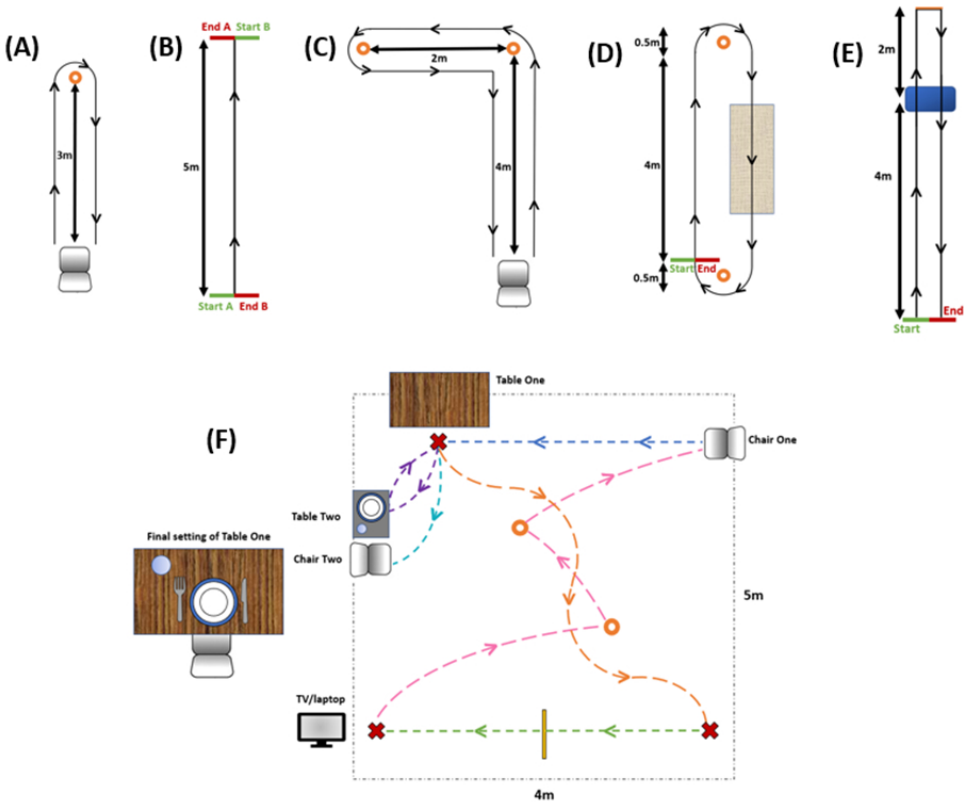


Illustration of the adopted marker set configuration.



The INDIP system.



Diagrams of the selected tasks: (A) Timed Up and Go, (B) Straight Walking Test, (C) L-Test, (D) Surface Test, (E) Hallway Test, F) Daily living activities. The tasks here are split into eight steps: (1) stand and walk to the red cross/cone next to the table, (2) set table one for dinner using the supplies on table two. (3) Sit down for a quick break and a drink (4) stand up and clear table one by placing all the items back on table two. (5) Walk to corner of the lab diagonally opposite while snaking around the cones in the middle of the floor. (6) Walk straight to the tv/laptop and (7) stand to watch a 1-minute video on the TV/laptop or have a conversation with the researcher. The participant will be advised that if they wish to sit before the minute is over a chair will be brought over by a researcher. (8) Walk back to chair one and sit down.

BMJ Open

A MULTI-CENTRIC OBSERVATIONAL STUDY FOR THE TECHNICAL VALIDATION OF REAL-WORLD MONITORING OF GAIT

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-050785.R1
Article Type:	Protocol
Date Submitted by the Author:	05-Aug-2021
Complete List of Authors:	Mazzà, Claudia; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Alcock, Lisa; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Aminian, Kamiar; Ecole Polytechnique Federale de Lausanne, Laboratory of Movement Analysis and Measurement Becker, Clemens; Robert Bosch Gesellschaft für Medizinische Forschung Bertuletti, Stefano; University of Sassari, Department of Biomedical Sciences Bonci, Tecla; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Brown, Philip; Newcastle Upon Tyne Hospitals NHS Foundation Trust Brozgol, Marina; Tel Aviv Sourasky Medical Center, Center for the Study of Movement, Cognition and Mobility, Neurological Institute Buckley, Ellen; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Carsin, Anne-Elie; Instituto de Salud Global Barcelona Caruso, Marco; Politecnico di Torino Dipartimento di Elettronica e Telecomunicazioni; Politecnico di Torino, PolitoBIOMed Lab – Biomedical Engineering Lab Caulfield, Brian; University College Dublin, Insight Centre for Data Analytics, O'Brien Science Centre; University College Dublin, UCD School of Public Health, Physiotherapy and Sports Science Cereatti, Andrea; Politecnico di Torino Dipartimento di Elettronica e Telecomunicazioni Chiari, Lorenzo; University of Bologna, Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi»; University of Bologna, Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV) Chynkiamis, Nikolaos; Northumbria University Newcastle, Department of Sport, Exercise and Rehabilitation Ciravegna, Fabio; The University of Sheffield, Department of Computer Science; The University of Sheffield, INSIGNEO Institute for in silico Medicine Del Din, Silvia; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences

	Eskofier, Björn; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering Evers, Jordi; McRoberts BV Garcia Aymerich, Judith; Instituto de Salud Global Barcelona Gazit, Eran; Tel Aviv Sourasky Medical Center, Center for the Study of Movement, Cognition and Mobility, Neurological Institute Hansen, Clint; University Medical Center Schleswig-Holstein Campus Kiel, Department of Neurology Hausdorff, Jeffrey; Tel Aviv Sourasky Medical Center, Center for the Study of Movement, Cognition and Mobility, Neurological Institute; Tel Aviv University, Department of Physical Therapy, Sackler Faculty of Medicine & Sagol School of Neuroscience Helbostad, Jorunn; Norwegian University of Science and Technology, Department of Neuromedicine and Movement Science Hiden, Hugo; University of Newcastle upon Tyne, School of Computing Hume, Emily; Northumbria University Newcastle, Department of Sport, Exercise and Rehabilitation Paraschiv-Ionescu, Anisoara; Ecole Polytechnique Federale de Lausanne, Laboratory of Movement Analysis and Measurement Ireson, Neil; The University of Sheffield, Department of Computer Science; The University of Sheffield, INSIGNEO Institute for in silico Medicine Keogh, Alison; University College Dublin, Insight Centre for Data Analytics, O'Brien Science Centre; University College Dublin, UCD School of Public Health, Physiotherapy and Sports Science Kirk, Cameron; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Kluge, Felix; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering Koch, Sarah; Instituto de Salud Global Barcelona Küderle, Arne; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering Lanfranchi, Vitaveska; The University of Sheffield, Department of Computer Science; The University of Sheffield, INSIGNEO Institute for in silico Medicine Maetzler, Walter; University Medical Center Schleswig-Holstein Campus Kiel, Department of Neurology Micó-Amigo, M. Encarna; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Mueller, Arne; Novartis Pharma AG, Novartis Institutes of Biomedical Research Neatrour, Isabel; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Niessen, Martjin; McRoberts BV Palmerini, Luca; University of Bologna, Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi»; University of Bologna, Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV) Pluimgraaff, Lucas; McRoberts BV Reggi, Luca; University of Bologna, Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV) Salis, Francesca; University of Sassari, Department of Biomedical Sciences Schwickert, Lars; Robert Bosch Gesellschaft für Medizinische Forschung Scott, Kirsty; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Sharrack, Basil; Sheffield Teaching Hospitals NHS Foundation Trust,

	Department of Neuroscience and Sheffield NIHR Translational Neuroscience BRC Sillen, Henrik; AstraZeneca Sweden, Digital Health R&D Singleton, David; University College Dublin, Insight Centre for Data Analytics, O'Brien Science Centre; University College Dublin, UCD School of Public Health, Physiotherapy and Sports Science Soltani, Abolfazi; Ecole Polytechnique Federale de Lausanne, Laboratory of Movement Analysis and Measurement Taraldsen, Kristin; Norwegian University of Science and Technology, Department of Neuromedicine and Movement Science Ullrich, Martin; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Computer Science Van Gelder, Linda; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Vereijken, Beatrix; Norwegian University of Science and Technology, Department of Neuromedicine and Movement Science Vogiatzis, Ioannis ; Northumbria University Newcastle, Department of Sport, Exercise and Rehabilitation Warmerdam, Elke; University Medical Center Schleswig-Holstein Campus Kiel, Department of Neurology Yarnall, Alison; Newcastle University Faculty of Medical Sciences, Translational and Clinical Research Institute; Newcastle Upon Tyne Hospitals NHS Foundation Trust Rochester, Lynn; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences; Newcastle Upon Tyne Hospitals NHS Foundation Trust
Primary Subject Heading:	Diagnostics
Secondary Subject Heading:	Research methods, Neurology, Respiratory medicine
Keywords:	Multiple sclerosis < NEUROLOGY, Parkinson-s disease < NEUROLOGY, Hip < ORTHOPAEDIC & TRAUMA SURGERY, Chronic airways disease < THORACIC MEDICINE, Heart failure < CARDIOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

A MULTI-CENTRIC OBSERVATIONAL STUDY FOR THE TECHNICAL VALIDATION OF REAL-WORLD MONITORING OF GAIT

Mazzà Claudia[1,2], Alcock Lisa[3], Aminian Kamiar[4], Becker Clemens[5], Bertuletti Stefano[6], Bonci Tecla[1,2], Brown Philip[7], Brozgol Marina[8], Buckley Ellen[1,2], Carsin Anne-Elie[9], Caruso Marco[10,11], Caulfield Brian[12,13], Cereatti Andrea[10], Chiari Lorenzo[14,15], Chynkiamis Nikolaos[16], Ciravegna Fabio[1,17], Del Din Silvia[3], Eskofier Björn[18], Evers Jordi[19], Garcia-Aymerich, Judith[9], Gazit Eran[8], Hansen Clint[20], Hausdorff Jeffrey[8,21], Helbostad Jorunn[22], Hiden Hugo[23], Hume Emily[16], Paraschiv-Ionescu Anisoara[4], Ireson Neil[1,17], Keogh Alison[12,13], Kirk Cameron[3], Kluge Felix[18], Koch Sarah[9], Küderle Arne[18], Lanfranchi Vitaveska[1,17], Maetzler Walter[20], Micó-Amigo M. Encarna[3], Mueller Arne[24], Neatrour Isabel[3], Niessen Martjin[19], Palmerini Luca[14,15], Pluimgraaff Lucas[19], Reggi Luca[15], Salis Francesca[6], Schwickert Lars[5], Scott Kirsty[1,2], Sharrack Basil[25], Sillen Henrik[26], Singleton David[12,13], Soltani Abolfazi[4], Taraldsen Kristin[22], Ullrich Martin[18], Van Gelder Linda[1,2], Vereijken Beatrix[22], Vogiatzis, Ioannis[16], Warmerdam Elke[20], Yarnall Alison[3,7], Rochester Lynn[3,7]

Corresponding author:
Prof. Claudia Mazzà
c.mazza@sheffield.ac.uk
Pam Liversidge Building - Room F14
Mappin Street, Sheffield, S1 3JD, UK
Phone: +44 (0)114 2226073

1. INSIGNEO Institute for in silico Medicine, The University of Sheffield, Sheffield, UK
2. Department of Mechanical Engineering, The University of Sheffield, Sheffield, UK
3. Translational and Clinical Research Institute, Faculty of Medical Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne, UK
4. Laboratory of Movement Analysis and Measurement, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland
5. Robert Bosch Gesellschaft für Medizinische Forschung, Stuttgart, Germany
6. Department of Biomedical Sciences, University of Sassari, Sassari, Italy
7. Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
8. Center for the Study of Movement, Cognition and Mobility, Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel
9. Instituto de Salud Global Barcelona, Barcelona, Spain
10. Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Turin, Italy
11. PolitoBIOMed Lab – Biomedical Engineering Lab, Politecnico di Torino, Turin, Italy
12. Insight Centre for Data Analytics, O'Brien Science Centre, University College Dublin, Dublin, Ireland
13. UCD School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland
14. Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi», University of Bologna, Bologna, Italy
15. Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV), University of Bologna, Bologna, Italy
16. Department of Sport, Exercise and Rehabilitation, Northumbria University Newcastle, Newcastle upon Tyne, UK
17. Department of Computer Science, The University of Sheffield, Sheffield, UK
18. Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

19. McRoberts BV, The Hague, Netherlands
20. Department of Neurology, University Medical Center Schleswig-Holstein Campus Kiel, Kiel, Germany
21. Department of Physical Therapy, Sackler Faculty of Medicine & Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel
22. Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway
23. School of Computing, University of Newcastle upon Tyne, Newcastle upon Tyne, UK
24. Novartis Institutes of Biomedical Research, Novartis Pharma AG, Basel, Switzerland
25. Department of Neuroscience and Sheffield NIHR Translational Neuroscience BRC, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
26. Digital Health R&D, AstraZeneca, Sweden

Keywords

Gait analysis; wearable sensors; activity monitoring; users evaluation; mobility outcomes

Wordcount: 5491

1 ABSTRACT

Introduction

Existing mobility endpoints based on functional performance, physical assessments and patient self-reporting are often affected by lack of sensitivity, limiting their utility in clinical practice. Wearable devices including inertial measurement units (IMUs) can overcome these limitations by quantifying digital mobility outcomes (DMOs) both during supervised structured assessments and in real-world conditions. The validity of IMU-based methods in the real-world, however, is still limited in patient populations. Rigorous validation procedures should cover the device metrological verification, the validation of the algorithms for the DMOs computation specifically for the population of interest and in daily life situations, and the users' perspective on the device.

Methods and analysis

This protocol was designed to establish the technical validity and patient acceptability of the approach used to quantify digital mobility in the real-world by Mobilise-D, an EU-funded consortium part of the Innovative Medicine Initiative, aiming at fostering regulatory approval and clinical adoption of DMOs. After defining the procedures for the metrological verification of an IMU based device, the experimental procedures for the validation of algorithms used to calculate the DMOs are presented. These include laboratory and real-world assessment in 120 participants from five groups: healthy older adults; chronic obstructive pulmonary disease, Parkinson's disease, multiple sclerosis, proximal femoral fracture, and congestive heart failure. DMOs extracted from the monitoring device will be compared to those from different reference systems, chosen according to the contexts of observation. Questionnaires and interviews will evaluate the users' perspective on the deployed technology and relevance of the mobility assessment.

Ethics and Dissemination

The study, registered at ISRCTN (12246987), has been granted ethics approval by the centre's committees (London – Bloomsbury Research Ethics committee; Helsinki Committee, Tel Aviv Sourasky Medical Centre; Medical Faculties of The University of Tübingen and of the University of Kiel). Data and algorithms will be made publicly available.

Strengths and limitations of this study

- A multi-disciplinary approach was implemented to define a protocol for the validation of tools for mobility monitoring, covering aspects related to devices, algorithms and users.
- A set of rigorous quality assurance procedures have been established to allow for the creation of a high-quality annotated dataset to foster development in the field of digital mobility monitoring.

- Mobility data will be collected in the laboratory and in the real-world for five different cohorts of slow-walkers. and subsequently will be made publicly available at the end of the study.
- The multi-stage and multi-device experimental procedures required by this validation study can be extremely challenging for both the participants and the assessors.
- For the laboratory acquisitions, the level of agreement between the gold standard and the inertial sensor devices might be affected by the limitations associated with a restricted capture volume.

2 INTRODUCTION

The ability to move is a key contributor to physical, mental and social well-being, which is in line with the World Health Organisation's (WHO's) definition of health [1]. However, the study of mobility has received relatively little attention, except for diseases characterised by specific mobility dysfunction. The increasing longevity of the world's population together with prolonged survival of many patients with long term conditions means that more people are suffering from loss of mobility, which in turn is a major contributing factor to a loss of independence [2,3]. This has a considerable and growing personal, societal and economic impact. Efforts to mitigate this loss of mobility are an increasing priority and promising interventions are now under investigation.

Existing mobility endpoints based on performance, patient self-reporting and one-off assessment are resource-intensive and lack sensitivity [4], which limits therapeutic development and clinical management. A novel approach is needed that is low cost, simple, accurate and that can be used in the real world, including the home and the community.

Poor gait, especially slow walking, is a key determinant of mobility loss. It is associated with greater mortality, morbidity, cognitive decline, dementia and fall risk [2,3]. Quantifying gait related mobility outcomes, including features such as step/stride duration and their variability, walking speed and asymmetry features is well established in supervised instrumented assessments.

Wearable devices including inertial measurement units (IMUs) that allow digital mobility outcomes (DMOs) to be described, are leading the transition from laboratory-based assessment of mobility (mobility capacity), to continuous, unsupervised monitoring of mobility in daily life conditions (mobility performance). Nonetheless, the validity of IMU-based methods to characterise real-world mobility, and gait in particular, is still limited, especially in populations suffering from pathological conditions. This is because measuring real-world gait is far from simple or straightforward. In addition, complex factors arise from multiple sources that influence outcome measures, including disease characteristics, patient specific habits, environment/context and the purpose of walking. All these factors limit the validity of existing algorithms developed to quantify targeted DMOs [5]. Additionally, validation should include simultaneous evaluation [6] of the participants perception and acceptability of the device [7] as well as aspects related to wearability and usability [8,9]. Finally, a separate assessment of the metrological performance of the sensors contained in the adopted device is required.

All these validation steps need to be taken before the DMOs and the associated technologies can be effectively used for clinical [4] and regulatory [10] purposes (see Figure 1 for summary).

[Figure 1 around here]

In this paper, we present the comprehensive, multi-stage protocol deployed in the technical validation study (TVS) that we have developed as part of the IMI2-JU-funded Mobilise-D project (Number 820820 [11]), that aims to validate a new digital method for remote monitoring of mobility. In particular, this multi-stage protocol aims to: a) verify the metrological performance of the sensors included in a IMU based monitoring device, using a procedure that could be replicated on any device; b) establish the validity and reliability [12] of the DMOs estimated by the algorithms using data from an IMU based device, taking into accounts the effects of populations (e.g., healthy adults, patients with various conditions), locomotor activities (simple straight walking versus complex walking tasks), contexts (lab based versus real-world), durations (device wearing time, DMOs hourly and daily fluctuations, etc.), and contextual confounding factors (such as location of walks, weather, use of walking aids, etc.); and, c) establish participants' and assessors' opinions on the usability and acceptability of the monitoring devices that will be deployed.

This protocol will provide the stakeholders, the first comprehensive, multi-modal solution to validate real-world mobility. Notably, the protocol described here has been accepted by the European Medical Agency [13] as part of the Mobilise-D process for the regulatory qualification of real-world mobility performance biomarkers in Parkinson's disease. [10]

3 METHODS

The monitoring device that will be used in this study to collect mobility data is the DynaPort MM+ IMU (McRoberts, Table 1). In a trade-off between usability, accuracy and ability to provide DMOs for both at step and stride level, this is attached to the lower back via an elastic waistband and Velcro strap. The validity of algorithms to accurately estimate DMOs in the real-world will be investigated in 120 participants including: healthy older adults (HA) and in five clinical cohorts (chronic obstructive pulmonary disease, COPD; Parkinson's disease, PD; multiple sclerosis, MS; proximal femoral fracture, PFF; and congestive heart failure, CHF), chosen as presenting a variety of gait and mobility features [4]. The usability and acceptability of the device from the perspective of the participants and the assessors involved in the study will be established via interviews and questionnaires.

3.1 Verification of the metrological performances of the device

Verifying the performance of a device needs a robust and comprehensive metrological characterisation of all the sensors that it embeds. This requires a series of standardised procedures (spot-checks) to be

implemented to ensure accuracy of raw data that will be used as input to the algorithms. Table 1 shows the sensing characteristics of the device used in this study, the Dynaport MM+ (dimensions: 106.6 x 58 x 11.5mm, size: 55 grams).

Table 1 Characteristics of the sensors included in the DynaPort MM+ device

Sensor	Sampling Frequency	Sensor Range	Sensor resolution
Tri-axial accelerometer	100 Hz	± 8g	1 mg (at ± 8g)
Tri-axial gyroscope	100 Hz	± 2000 dps	70 mdps (at ± 2000 dps)

According to Institute of Electrical and Electronics Engineers *Standard for Sensor Performance Parameter Definitions ((IEEE 2700-2017* [14]) a number of parameters are needed to characterise the metrological performance of the accelerometer, gyroscope, and magnetometer included in an IMU. These parameters can be computed under both static and dynamic conditions. The main noise parameters used are the first (mean value) and second (variance) order statistics, and the root Allan variance parameters of noise [15,16]. The parameters related to the first and second order statistics of noise can also be estimated by means of a short static acquisition (the minimum length of each acquisition is defined by IEEE standard for each sensor). These short static acquisitions can be performed simply using a plastic cube, where the device can be properly secured and then each of the three sensors' coordinate axes x, y and z are in turn aligned with the direction of gravity (g) as well as its opposite direction (six combinations, Figure 2). The root Allan variance parameters are instead computed over a long static acquisition [17]. Typically, the acquisitions are performed over a period of 4-8 hours [15,16]. All the static acquisitions should be carried out at a constant temperature of 25°C.

In comparison to a static acquisition, characterising the dynamic metrological performance of the sensors embedded in an IMU is less straightforward, since the metrological standards provided by IEEE describe a sequence of operations requiring an expensive and complex testing instrumentation. However, various alternatives have been proposed in the literature. The accuracy of a gyroscope can be quantified during a single-axis rotation by a known angle by computing the ideal angular velocity and comparing it to the average measured angular velocity [18,19]. This procedure should be performed using a rotation plate with a rotating speed comparable to what is encountered during human gait (~200 deg/s) (Figure 2).

[Figure 2 around here]

The tests described above will be performed on thirty-five different DynaPort MM+ devices deployed in the study. This will allow conformity with manufacturer indications to be verified, highlight the need for sensor recalibrations and provide benchmark standards for any device with equivalent sensing capacity. In turn, this will allow any device with an equivalent or superior solution to be used, in order

to facilitate broader adoption of validated algorithms and cope with a continuously changing hardware landscape.

3.2 Protocol for Validation of the algorithms

Several algorithms to detect the DMOs from a single device have been implemented according to agreed definitions [20] and based on existing literature. At this time, these algorithms are being concurrently validated using the approach described by Bonci et al. [6] using pre-existing datasets, which mostly include lab-based observations only. Following this selection process, the best performing algorithms will be assessed using the data captured with the protocol here described.

3.2.1 Ethics and Dissemination

This multi-centre study is sponsored and coordinated by The Newcastle upon Tyne Hospitals NHS Foundation Trust, UK. Participants will be recruited in five sites across Europe: Tel Aviv Sourasky Medical Center, Israel (ethics approval granted by the Helsinki Committee, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel, 0551-19TLV), Robert Bosch Foundation for Medical Research, Germany (ethics approval granted by the ethical committee of the medical faculty of The University of Tübingen, 647/2019BO2), University of Kiel, Germany (ethics approval granted by the ethical committee of the medical faculty of Kiel University, D438/18), The Newcastle upon Tyne Hospitals NHS Foundation Trust, UK and Sheffield Teaching Hospitals NHS Foundation Trust, UK (ethics approval granted by London – Bloomsbury Research Ethics committee, 19/LO/1507).

As per the study register (ISRCTN, 12246987) the data collection was originally planned to start in April 2021 and last six months. The pandemic situation meant that the study started in July 2020 and is now planned to finish in September 2021. All experimental procedures are constantly monitored and revised as needed to ensure full compliance with COVID19-related health and safety measures and for safeguarding of both the study participants and the assessors. The data collected as part of this protocol will be made publicly available together with the algorithms used to process the data.

3.2.2 Participants

A convenience sample of 120 participants will be recruited via their clinical care team or research registers to represent the five disease cohorts (COPD, PD, MS, PFF, and CHF), as well as healthy older adults (HA). Twenty participants will be recruited for each cohort. Each cohort will be recruited across multiple sites to ensure generalisability (e.g. differing cultures and contexts). Inclusion and exclusion criteria, grouped by total cohort and disease cohort, are summarised in Table 2. Given the novelty of the data, rather than on a power calculation the sample size of 120 has been initially defined according to Consensus-based Standards for the selection of health Measurement Instruments guidelines for measurement properties (COSMIN [21]). This sample size, however, will be refined after 50% of the

data collection. Given the DMOs are measured at walking bout level and not at patient level, in this analysis we will use the effective number of walking bouts observed during the 2.5 hours to perform the power calculation. We will base this analysis on a desired ICC coefficient ≥0.7, with Alpha=0.05 and Beta=0.9, and an aimed confidence interval of 0.1. Based on this review, more participants may be recruited.

All participants will give written informed consent prior to undergoing a clinic/laboratory-based session to record generic and disease-specific characteristics. This will include participant reported outcomes, assessments and medical notes review. The generic and cohort-specific clinical outcomes that will be collected are summarised in Table 3.

Table 2 Inclusion and exclusion criteria adopted for the different disease cohorts

Group	Inclusion criteria	Exclusion criteria
All groups	-able to walk 4 meters independently with or without walking aids -able to give informed consent -willingness to wear the sensor set-ups during the study -shoe size 36 EU (3 UK) or above -able to read and write in first language of the respective country -Montreal Cognitive Assessment (MoCA) >15 [22] -available for home /office visit during study period	-occurrence of any of the following 3 months prior to inclusion: myocardial infarction, hospitalisation for unstable angina, stroke, coronary artery bypass graft, percutaneous coronary intervention, implantation of a cardiac resynchronisation therapy device -current medical condition that could interfere with the patient's compliance
COPD	- ≥45 years of age -Diagnosis of COPD (post-bronchodilator forced expiratory volume in the first second (FEV_1) to forced vital capacity (FVC) ratio <0.70) -clinical stability, defined as at least 4 weeks without antibiotics and/or oral corticosteroids to treat either a moderate or severe exacerbation - current or ex-smokers with a smoking history equivalent to at least 10 pack years (1 pack year = 20 cigarettes smoked per day for 1 year)	-having undergone major lung surgery (e.g., lung volume reduction, lung transplant) -having a lung tumour -primary respiratory diseases other than COPD (e.g., asthma) - impaired mobility related to non-COPD causes, as judged by the investigator
PD	-aged 18+ years -Diagnosis of PD according to the Movement Disorders Society criteria [23]	-impaired mobility related to non-PD causes, as judged by the investigator
MS	-aged 18+ years -Diagnosis of MS based on the revised McDonald's criteria	-impaired mobility related to non-MS causes, as judged by the investigator

PFF	-65+ years of age -surgical treatment (fixation or arthroplasty) for a low-energy fracture of the proximal femur (ICD*-10 diagnosis S72.0, S72.1, S72.2) as diagnosed on X-rays of the hip and pelvis within last 12 months	-impaired mobility related to non-PFF causes, as judged by the investigator
CHF	- ≥45 years of age -Diagnosis of chronic heart failure with a grading of II-IV of the New York Heart Association (NYHA) Classification	- history of COPD ≥GOLD III - impaired mobility related to non-CHF causes, as judged by the investigator
HA	65+ years of age	
	*ICD, International Classification of Disease	

Table 3 Generic and cohort-specific clinical outcomes.

Cohort	**Generic Outcomes**
All	- Descriptive measures (age, sex, living arrangements, education) - Anthropometric measures (height, mass, shoe size, waist width) - Health status (comorbidities, number of falls and injuries in the 12 months prior to assessment, walking aid usage and current medication) - Montreal Cognitive Assessment to evaluate global cognition [22] - Visual Analogue Scale to measure pain during walking (0-10, from no pain to worse pain possible) - Function component of the Late-Life Function and Disability Instrument (LLFDI) to evaluate function and disability [24,25]

Cohort	**Cohort-specific clinical outcomes**
PD	- Movement Disorder Society (MDS)-sponsored revision of the Unified Parkinson's Disease (PD) Rating Scale (UPDRS), motor part [26]
MS	- Expanded Disability Status Scale (EDSS)* [27]
COPD	- 6-minute walk test - Recent spirometry test obtained from medical notes to characterise lung function* - COPD Assessment Test™ (CAT) [28]
PFF	- Short Physical Performance Battery (SPPB) – quiet standing balance task, a five times chair-raise test, and a 4m walk test at preferred gait speed
CHF	- 6-minute walk test - Kansas City Cardiomyopathy Questionnaire (KCCQ) * [29]

* denotes measures obtained from medical records if completed within 6 months prior to assessment

3.2.3 Patient and Public Involvement

Patient and Public involvement and engagement (PPIE) has informed the design and conduct of this study. The protocol and patient facing documents were reviewed and changes implemented based upon reviewer feedback. We will work with our Patient Advisory Group (with members representing the patient cohorts involved in the study) to review study findings, data interpretation and study reporting and dissemination, including co-design and presentation of dissemination materials for patients and the public. The learnings derived from these activities will further inform the work of the wider Mobilise-D project.

3.2.4 Experimental protocol

Performance of algorithms to determine walking-related DMOs is mostly affected by three factors: (i) the type of motor task (e.g., slow as opposed to fast, straight as opposed to curvilinear or inclined walking, etc.), (ii) the population of interest (e.g., healthy vs pathological gait), (iii) the context of observation (e.g., home vs outdoors). To accommodate these factors, we have developed a comprehensive, multi-stage protocol that includes a variety of tests conducted both in a laboratory context and in the real world (Table 4).

Table 4 Summary of the experimental protocol used for the validation of the algorithms in the laboratory and in the real world. INDIP= INertial module with DIstance Sensors and Pressure insoles.

Context of assessment	Reference Systems	Tested device	Mobility Tasks
Laboratory	Stereophotogrammetry	DynaPort MM+	Structured mobility tasks and daily living activities
		INDIP	
Real world (2.5 hours)	INDIP	DynaPort MM+	Unsupervised real-world activities (including predetermined tasks)
	Mobile Phone with Aeqora App		
	Beacon		
Real world (7 days)	Mobile Phone with Aeqora App	DynaPort MM+	Unsupervised daily living
	Beacon		

a. Laboratory based assessment

Laboratory-based observations will be used to quantify validity and consistency within and between groups and different types of walking tasks under controlled ideal conditions. Structured and task-based mobility activities and a simulated daily activity session, mimicking habitual movements performed at home or at work will be included. The outcome of this comparison will provide the level of highest expected accuracy and minimum detectable changes for a given DMO.

Measurement tools

Reference system

A stereophotogrammetric (SP) system (100Hz) will be used as the gold standard in structured and simulated tests of daily activities to validate the DMOs calculated from the DynaPort MM+ raw data. SP systems provide a measurement of the instantaneous position of points in a 3D measurement volume, by means of a set of cameras, each of which can capture the 2D trajectories of markers that are attached to the object of interest. The trajectory reconstructions are affected by systematic and random instrumental errors, normally minimised to a few millimetres via ad hoc calibration procedures and filtering and smoothing techniques [30].

To ensure quality and consistency in the SP data collection, accommodating different SP systems across sites, a spot check designed following the methodology proposed by Di Marco et al. [31] will be used, which will establish the specific level of accuracy for each system. A graphical user interface (GUI) for automated pre-processing of the SP data will ensure consistency in associated procedures (labelling, gap filling, etc.).

A bespoke marker set will be adopted, including four markers on each foot for detecting the gait events and four markers on the lower back device to track the displacement of the DynaPort device (Figure 3).

[Figure 3 around here]

Tested devices

During this observation each participant will also be equipped with an additional multi-sensor system (INertial module with DIstance Sensors and Pressure insoles, INDIP) [32–34] and with the Dynaport MM+. The INDIP system (Figure 4) includes four inertial modules (one on the lower back, one on the non-dominant wrist and two on the feet), two distance sensors and two force-sensitive resistor pressure insoles including 16 force-resistive sensing elements (manufacturer 221e S.r.l., Italy). The INDIP has been designed to be used as a reference for real-world experiments, and in this phase of the protocol its performance will be validated against the SP system for the populations of interest. Spatio-temporal parameters will be estimated exploiting the sensors redundancy and implementing previously validated sensor fusion algorithms. Gait events will be detected using data from pressure insoles and inertial

sensors independently, and then combined to increase robustness and detection accuracy (missed and minimisation of extra events). Spatial variables will be computed from the inertial data of the feet using a Madgwick filter [35,36] combined with a zero-velocity update [37,38], subsequently velocity and displacement will be calculated using a direct and reverse approach [32,39–41]. The individual components of the INDIP system and the associated algorithms for the estimates of the DMOs have already been extensively validated in previous studies on various healthy and pathological cohorts [39,42,43]. The final assembled system in its fully synchronized configuration, developed to address the requirements of this study, is expected to perform equivalently and as such we can anticipate mean absolute percentage errors of 1% on the stride duration, between 2% and 3% in the estimate of the stride length. Preliminary results from in-lab validation showed percentage errors of about 2% for gait speed as estimated during continuous walking, including both straight and curvilinear portions [44].

[Figure 4 around here]

The lower back INDIP unit and the DynaPort MM+ will be rigidly attached to each other. The data from the SP (100Hz), INDIP (IMU and insoles, 100Hz, Distance sensor, 50Hz) and DynaPort MM+ (100Hz) systems will be synchronised using a hardware-based approach for the SP and the INDIP system, and timestamps to align recordings from the INDIP and the DynaPort MM+.

Mobility tasks

Structured Mobility Tasks

Straight Walking: Straight walking is the most common test of walking. [45,46] The participant walks for a distance of 5m from a standing start and will be repeated at three different walking speeds: preferred, fast and slow (Figure 5 A).

Timed Up and Go (TUG): The TUG is a widely used clinical assessment of a person's mobility [47]. The participant is asked to sit in a chair, stand up, walk 3 m in a straight line, make a 180° turn, walk back to the chair, turn and sit down (Figure 5 B).

L Test: The participant is asked to sit in a chair, stand up, walk straight, turn 90° to the left around a cone, walk straight to the second cone, make a 180° turn to the left, walk straight before making a final 90° turn to the right and return to the chair to sit down (Figure 5 C). Besides being a clinically validated test, [48] the main purpose of including this test is the variation in curvilinear walking and the inclusion of different types of turns.

Two novel additional tests were also included to simulate confounding factors that could be encountered in the real word:

Surface Test: The participant walks around a defined circuit by turning around the cones (Figure 5 D). The circuit is completed twice, creating the longest walking bout out of all the tasks (approx. 20m).

Hallway Test: The participant walks along a 6m walkway stepping up and down a step positioned in the walkway. At the end of the walkway, the participant will complete a sharp 180° turn and walk back

along the walkway (again stepping up and down off the step) until reaching the end point of the test (Figure 5 E).

[Figure 5 around here]

Daily living activities

These lab-based tasks will be used to simulate daily activities expected in the real life, similar to previous studies [49]. The participant starts by sitting in chair one and then executes a series of daily living tasks while moving around the room (see Figure 5F and G).

Patients will be given regular opportunities for rest periods and will be asked to communicate if they require any additional breaks or would like to stop the assessment at any point. Use of arm rests for the TUG, L-Test and SDA, as well as handrails for the hallway test are permitted when needed.

b. Real - world validation (2.5 hours observation)

This phase of the protocol will quantify validity and consistency across individuals and different types of walking tasks in the real world. It will be performed in a habitual environment (home/work/community/outdoor) chosen by the participants, without specific restrictions. The duration of the observation has been established as a trade-off between experimental, clinical and technical requirements.

Measurement tools

Reference systems

Participants will be asked to wear the INDIP, which in this phase of the protocol will be used as a reference system for the quantification of the DMOs provided by the single sensor algorithms, as applied to the DynaPort MM+ data.

In order to quantify the effects of contextual confounding factors, the participants will also be provided with a system detecting outdoors walking, gradient of descent/ascent (walking uphill/downhill). The system is developed as a mobile Android application (Aeqora app) and the device selected was a Samsung S9 with Android 10. The app is composed of three parts: *(i)* the core tracker, *(ii)* the interface and *(iii)* the server infrastructure collecting data across users. The core tracker, adapted from a library developed by the University of Sheffield [50], utilises the mobile phone's internal sensors to compute the type of activity (e.g., walking) and intensity (e.g., cadence) to identify geo-located bouts of movement. It operates in the background and senses mobility features through a range of sensors (e.g., step counters, activity recognition, accelerometer, gyroscope, etc.) as well as from location services (GPS, network, Bluetooth, etc.). It collects the data and stores the raw sensor data into a local database in real time. A set of mechanisms have been developed to control access to this data, keep it secure, and regulate its use. First, no user identity information is sent within a single request as a token identifier is

used. Additionally, a security layer is built based on Secure Socket Layer (SSL) and Transport Layer Security (TLS) 3.0 protocol to the data with scalable and efficient encryption algorithms. An SSL/TLS certificate is issued and used to establish identity and trust between server and client apps (desktop and mobile), ensuring privacy and security whenever communicating sensitive data.

Data collected during the experiments will be sent to a cluster of servers that uses algorithms to integrate the phone's data with contextual information about the locations where the participant will walk: where possible walks will be matched to OpenStreetMap [51] roads and paths, to remove GPS noise, the slope variation of each walk is computed on tiles, with a resolution of 5m within the UK (using Ordnance Survey Terrain 5) [52] and 30m in the other locations (using NASA's Shuttle Radar Topography Mission (SRTM) data [53], indoors and outdoors walking is recognised. Moreover, weather is associated with participant location based on the most proximate weather station.

The use of walking aids will also be monitored in this phase. For this purpose, a Bluetooth beacon (BlueBeacon Tag, BlueUp) will be attached to the walking aid and its activity will be detected by the phone's mobile tracker and saved by the app. The distance between the phone and the Beacon and data from the accelerometer contained in the Beacon will be integrated to determine when the aid is in use.

The above contextual factors and the use of walking aids will be included in the analyses to determine the extent to which they affect variation in the DMOs, although the degree of correlation will be adversely affected by the issues in accurately measuring context that are associated with missing data and GPS accuracy.

Mobility Tasks

To capture the largest possible range of activities during this assessment, participants will be guided by the following list of activities to be included: if relevant for their chosen environment, rise from a chair and walk to another room; walk to the kitchen and make a drink; walk up and down a set of stairs (if possible); walk outdoors (if possible, for a minimum of two minutes); if walking outside, walk up and down an inclined path. No supervision or structure to how these tasks should be completed will be given to the participants.

c. Real - world validation: seven days monitoring

This observation will quantify the effects of device wearing time, hourly and daily fluctuations of DMO's, and contextual confounding factors (such as location of the walk, weather, type of housing, etc.).

Measurement tools

The participants will be asked to wear the DynaPort MM+, and to carry a mobile phone equipped with the Aequora App. Bluetooth beacons will also be used to track the use of walking aids. The participants will wear the Dynaport MM+ at all times (including at night). As this device is not waterproof, they

will be instructed to remove it for showering, bathing, using a sauna and swimming and reattach it afterwards. They will be asked to keep the mobile phone charged, switched on at all times and to carry it with them whenever possible, especially when leaving the house.

Mobility Tasks

Participants will be monitored continuously for seven days, without any specific instruction being provided, except for that of wearing the provided measurement tools.

3.3 Assessment of participants' and assessors' experience

This part of the study will evaluate the participants' and assessors' experience of using the monitoring device. For the participant's assessment, wear-time of the device during the 7 days monitoring will be collected as a primary measure of compliance. Following the period of the seven-day, free-living data collection, participants will complete two questionnaires to assess the acceptability of the device. The first is a 12-item questionnaire [54] investigating usability on a 5-point ordinal scale. The questions are simple and focus on the impact of using a wearable device on participants' feelings, comfort and the ease of use of the device. The second questionnaire is the Comfort Rating Scale, [55] a 6-item measure investigating the comfort of a wearable device on a 21-point ordinal scale from '0 – low agreement' to '20 – high agreement'.

A subset of participants from all recruiting sites and cohorts will complete a semi-structured interview (see supplementary file). For this qualitative part of the study, sampling will continue until saturation, i.e. until no additional learning is identified from the data. The interview will explore participants' opinions on the use of wearable devices and digital technology in healthcare, experiences of managing their condition, experiences of technology, and opinions on data privacy associated with the use of technology in healthcare. Additionally, participants will be asked about their experiences of using the device, including comfort, perceived usefulness and ease of use, barriers and facilitators, and any other usability experiences that they may have encountered. All interviews will be audio-recorded, transcribed verbatim and, where required, translated to English.

To assess the professionals' experience, assessors from each of the clinical sites will be asked to assess the usability of the device after completion of the data collection. They will be provided with three questionnaires: 1) The System Usability Scale [56] a commonly used, validated 10-item questionnaire that asks users to rate a device on a 5-point Likert scale from '1 strongly disagree' to '5 – strongly agree'. Questions focus on the ease of use of the device, and the integration of various functions within it; 2) The IBM Computer System Usability Questionnaire [57] (to assess the DynaPort MM+ software), is composed of 19 items and asks respondents to consider their interaction with a computer system on a 7-point Likert scale from the perspective of data collection; 3) A bespoke questionnaire designed specifically for the TVS to assess the acceptability and effectiveness of the training methods, procedures, and any other materials provided within the study. The questionnaire will ask respondents to rate their experiences on a 7-point Likert scale to determine whether any changes to the procedures

and materials are required, and whether training was effective in preparing assessors to implement the assessment protocol as planned.

In addition, assessors will complete a semi-structured interview with the aim of exploring their experiences of the data collection process. Assessors will be asked about the use of the device (e.g., ease of use, intuitiveness, data collection and download procedures, etc.), training and materials provided prior to the commencement of the study, and barriers and facilitators to using the device. The topic guide and open-ended questions allow for new areas of conversation to emerge. All interviews will be audio-recorded and transcribed verbatim.

3.4 Data management

All data will be uploaded to a central platform "e-Science Central" (e-SC) [58] which provides data processing and storage functionality in accordance with principles of reproducible research. The underlying infrastructure complies with ISO 27000 standards for Information Security Management Systems and is hosted on Amazon Web Service secure services cloud platform. Data will be integrated on the platform by means of implementation of a standardised file nomenclature system. At point of capture, each file will be labelled in standardised format. For source data, we will adhere to principles defined by the US Food and Drug Administration [59] for making them attributable, legible, contemporaneous, original, and accurate (ALCOA+). In particular, we will use both web-based forms on e-SC and an application from ERT (partner in the project) to capture the electronic clinical outcome assessments (eCOAs). The e-SC forms provide storage of event data, and support for data validation and basic data entry and verification. Both e-SC and ERT systems employ error-handling at source which alert the assessors of incorrect data entry (e.g., min/max boundaries, required/optional fields). Data captured at source on paper will be copied, signed, and scanned, then uploaded to e-SC as a certified copy. The motion capture data will also be transferred to e-SC and stored in an unmodified form. These data will be either uploaded directly to e-SC via the e-SC portal or transferred via an Application Programming Interface (API). The algorithms being developed and benchmarked will be used to process these files and extract and store the DMOs.

3.5 Data analysis plan

3.5.1 Verification of the device

Mean and standard deviation readings of the accelerometer, gyroscope, and magnetometer signals captured during static acquisition will be used to assess the reliability of the manufacturer sensor calibration and to detect the presence of abnormal spikes in the sensor signals. Data from long static acquisitions will be used to confirm the stability of sampling frequency, the duration of the battery, and to estimate the Allan deviation (bias instability) of the gyroscope readings over time. Errors of

gyroscope readings will be assessed using mean and standard deviation of nominal, measured and relevant errors for angular velocity values during the dynamic acquisitions.

3.5.2 Validation of the algorithms

The data analysis will determine criterion validity (including selected performance metrics and criterion (concurrent) validity metrics of the primary (real world walking speed, RWS) and secondary DMOs listed in Table 5. Table 6 summarises the statistical tools that will be used to quantify each of the DMOs. All statistical analyses will be performed using the statistical analysis toolbox of Matlab R2018a. In all tasks and observations, continuous variables (e.g., cadence, real-walking-speed) will be summarised with descriptive statistics for the values obtained within walking bouts (mean and standard deviation). In addition, the mean, minimum, maximum, standard deviation, median, interquartile range (IQR) and Root Mean Square Error (RMSE) of DMOs over all available walking bouts will be presented. Categorical variables (e.g., laterality of initial contacts) will be summarised with frequency counts and percentages. Confidence Interval 95% (CI) will be provided for the Interclass Correlation Coefficients (ICC).

Using the gold standard as a reference, True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN) will be identified for the DMOs identified from the single device using a cut-off tolerance window defined as a fixed interval of 0.5 seconds [60] and centred on each event detected by the reference system. The following performance metrics will then be calculated:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Positive\ Predicted\ Value = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Specificity = \frac{TN}{TN + FP} \text{ and}$$

$$F1\ score = 2 * \frac{Positive\ Predicted\ Value * Sensitivity}{Positive\ Predicted\ Value + Sensitivity}$$

Criterion validity will be characterised by evaluating the absolute and relative errors, defined as the relative and absolute differences between the DMOs quantified with the single device and those derived from the reference systems:

$$Relative\ error = \left(\left|\frac{DMO\ estimated\ by\ IMU - DMO\ estimated\ by\ Reference\ System}{DMO\ estimated\ by\ RS}\right|\right) \times 100$$

$$Absolute\ Error\ (DMO) = |DMO\ estimated\ by\ IMU - DMO\ estimated\ by\ Reference\ System|$$

The mean, standard deviation and maximum of all errors will be reported for each walking bout. Limits of agreement between single sensor and reference system DMOs will be quantified. In addition, statistically significant differences between the DMOs quantified by the IMU and those by the reference system, parametric (paired t-test) or non-parametric (Wilcoxon signed-rank test) tests will be performed depending on the normality of the distribution of the DMOs. Data distribution will be visually inspected with histograms, and normality tested with the Shapiro-Wilk test.

Concurrent validity between the DMOs quantified by the single device and those derived from the reference systems will be evaluated by quantifying the Intraclass Correlation Coefficient (ICC (2,1)).

All results will be presented separately by cohort (e.g., PD) and subgroup (i.e., subgroups of the cohorts) stratified by average stride gait speed (e.g., fast speed: walking speed > 1 m/s, medium speed: walking speed between 0.5m/s and 1 m/s, slow speed: walking speed < 0.5 m/s) [2].

If participants missed data from one assessment (e.g., one of the tasks in the laboratory) or observation (e.g., 2.5hs), their remaining available data will still be included in the analyses. Assuming that data are missing completely at random, a complete case approach will be used to handle missing data [61].

Table 5 List of digital mobility outcomes (primary and secondary DMOs) that will be analysed as part of the TVS.

Variables	DMOs (units)	Definition	DMO Attainable			
			DynaPort MM+	SP System	INDIP	Aeqora App
Walking Bout (WB) *A walking sequence containing at least two consecutive strides of both feet. Start and end of a WB are determined by a resting period or any other activity (non-walking period).*	Number of WBs (count)	Based on the identification of gait as an activity (yes/no) to a sample level of 0.1 s	✓	✓	✓	✓
	WB Start (s)	Start of WB	✓	✓	✓	
	WB End (s)	End of WB	✓	✓	✓	
	WB Duration (s)	Time between start and the end of WB	✓	✓	✓	
Stride/ Step Duration (SD) *Refers to the duration (time intervals) of strides, calculated as the time in between two non-consecutive (alternate) initial contacts.*	Stride Duration (s)	Duration between two non-consecutive (alternate) initial contact events	✓	✓	✓	
	Step Duration (s)	Duration between two consecutive initial contact events	✓	✓	✓	
Cadence (CE)	Cadence (steps/minute)	Steps performed within a minute	✓	✓	✓	
Stride Length (SL)	Mean Stride Length (m)	Average stride length within a WB	✓	✓	✓	
Real World Walking Speed (RWS)	Walking speed (m/s)	Velocity, average stride speed within a WB	✓	✓	✓	✓
Turning	Number of Turns	Overall number of turns performed in a WB based on the identification of turns (yes/no) to a sample level of 0.1 s	✓	✓	✓	
	Turn Start (s)	Start of each turn within the WB	✓	✓	✓	
	Turn End (s)	End of each turn within the WB	✓	✓	✓	
	Turn Duration (s)	Time between the start and the end of the turns within the WB	✓	✓	✓	
	Maximal Turn Angle (deg)	Maximal angle achieved in the turn	✓	✓	✓	

Height Estimation	**Elevation Change (m)**	**Difference between the minimal and maximal height or elevation for the complete walking bout detected for incline walking**	✓	✓	✓	✓
Left/Right Identification	**Laterality (label)**	**Left or Right category, indicating the foot with which the initial contact is performed**	✓	✓	✓	
Secondary Outcomes (SO)	**Number of Final Contact Events (counts)**	**Correct identification of Final Contact events**	✓	✓	✓	
	Final Contact Event (s)	**Instant of time at which each final contact event is performed within a walking bout**	✓	✓	✓	
	Swing Phase Duration (s)	**Time between the last contact of the current footfall and the first contact of the next footfall on the same foot**	✓	✓	✓	
	Stance Phase Duration (s)	**Time in between the first contact and the last contact of two consecutive footfalls on the same foot**	✓	✓	✓	
	Variability of: Step Time, Stride Time, Swing Time, Stance Time Stride Velocity Stride Length (same units as variable)	**St. Dev. and Coefficient of Variation of step time, of stride time, of swing time, of stance time, of stride velocity and stride length within a WB**	✓	✓	✓	
	Asymmetry of: Step Time, Stride Time, Swing Time, Stance Time (same units as variable)	**Asymmetry evaluated as difference between right and left steps or strides for step time, of stride time, of swing time and of stance time within a WB**	✓	✓	✓	
Contextual Factors	**Location (label indoor/outdoor)**	**WB completed in an indoor or outdoor environment**	✓			✓
	Walking Aid (label yes/no)	**Walking aid assistance during WB**	✓			✓

Table 6. List of statistical analyses and performance metrics that will be used for the various digital mobility outcomes (DMOs). Performance metrics and criterion validity are those that will be used to compare DMOs obtained from a single device versus those obtained from the reference system. The types of plots listed in the table will be used to visualise performance metrics and to support interpretation of the results.

DMO	PERFORMANCE METRIC					CRITERION VALIDITY							PLOTS		
	Sensitivity	Positive Predictive Value	Accuracy	Specificity	F1-score	Error (absolute & relative)	SD & Max of Error	Root Mean Squared Error	Precision	Concurrent validity (ICC)	Significant Difference	Limit of Agreement	Bland Altman Plots	Scatter Plots for Correlation	Histogram Plots
Number of Walking Bouts	✓	✓	✓	✓	✓										
WB Start						✓	✓	✓	✓				✓	✓	✓
WB End						✓	✓	✓	✓				✓	✓	✓
WB Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Initial Contact Events	✓	✓			✓	✓									
Step Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Stride Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Final Contact Events	✓	✓			✓	✓									
Mean Stride Length						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cadence						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Elevation Change						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Walking Speed						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Number of Turns	✓	✓	✓	✓	✓										
Turn Start						✓	✓	✓	✓				✓	✓	✓
Turn End						✓	✓	✓	✓				✓	✓	✓
Turn Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Maximal Turn Angle						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Laterality						✓	✓	✓	✓	✓					
Sequence Order						✓	✓	✓							

3.5.3 Participants' and assessors' experience

Participant and professional questionnaire data will be analysed using descriptive statistics. Interviews will be analysed using thematic analysis. [62] Transcripts will be deductively examined for the presence of themes related to the acceptability of the monitoring device to participants (i.e., comfort, interference with daily living) based on previous literature. Specifically, perceived usefulness, comfort, and ease of use are critical factors of usability thus, these will be the categories examined within the transcripts. Regarding participants use of technology to manage their healthcare condition an inductive approach will be taken. A list of codes relevant to the question will be generated and then refined by grouping them into potential themes.

Authors' contributions:

This study is part of a large collaborative initiative, evolving around the design and execution of the study here described. The complexity of the protocol and the highly multi-disciplinary content of the

study justifies the high number of authors that have been and will be involved in the different stages of its planning, conducting and in the design of the analysis plan.

Manuscript initial drafting: CM, TB, SDD, AC, LA, AK, KS, SB, FC, LR. Design and deployment of experimental tools and protocols: KS, LS, CM, TB, EB, FC, NI, BS, LV, IV, EH, NC, FS. AC, MC, SB, LRO, AY, CK, EA, LS, CB, WM, CH, EW, JHA, EG, MB; J, MN, KT, BV, JHE, AK, DS, AM, VL. Data collection: LA, TB, PB, MB, EB, CF, EG, CH, EH, WM, CB, LP, LS, KS, BS, LVG, IV, AY, EW. Data analysis and algorithm development: SDD, EA, HS, AS, BE, FK, AK, MU, AC, JG, SK, LC, LP, LR, HH, EG. Draft revisions: LRO, KA, BE, JHA, BV, HS, AM, JHO, PB, FK, BC, AK, LC, LP, EB, KS, SK, LP, HH, WM, NC, BS, AI, IN, VL, SDD. Study design and coordination: CM, LRO, AM, SDD, AC. All authors have read and approved the final manuscript.

Competing interests statement

Fabio Ciravegna is CEO and shareholder of Aeqora Ltd, of which Vitaveska Lanfranchi is director and shareholder. Henrik Sillén is an employee of AstraZeneca. Bjoern M Eskofier is co-founder and owns shares of Portabiles HealthCare Technologies GmbH. McRoberts is the manufacturer of the DynaPort. Martijn Niessen, Jordi Evers, and Lucas Pluimgraaff are employees of McRoberts. Arne Mueller is an employee of Novartis and holds stock in Novartis. Lars Schwickert and Clemens Becker are consultants of Philipps Healthcare, Bosch Healthcare, Eli Lilly, Gait-up. Jeffrey M. Hausdorff reports having submitted a patent for assessment of mobility using wearable sensors in 400 Parkinson's disease; the intellectual property rights 401 are held by the Tel Aviv Medical Center. Luca Palmerini and Lorenzo Chiari are co-founders and own shares of mHealth Technologies.

Funding

This study was co-funded by the European Union's Horizon 2020 research and innovation programme and EFPIA via the Innovative Medicine Initiative 2 (Mobilise-D project, grant number IMI22017-13-7-820820). The views expressed are those of the authors and not necessarily those of the IMI, the European Union, the EFPIA, or any Associated Partners. We acknowledge the support of Grünenthal GmbH via the funding of a PhD scholarship directly dedicated to the technical validation protocol.

We acknowledge support from the Spanish Ministry of Science and Innovation through the "Centro de Excelencia Severo Ochoa 2019-2023" Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program.

SDD, AJY and LR are also supported by the Newcastle Biomedical Research Centre (BRC) based at Newcastle upon Tyne and Newcastle University. CM, BS, LVG and EB are also supported by the Sheffield Biomedical Research Centre (BRC) based at the Sheffield Teaching Hospital and the University of Sheffield. The work was also supported by the NIHR/Wellcome Trust Clinical Research Facility (CRF) infrastructure at Newcastle upon Tyne Hospitals NHS Foundation Trust and the CRF at the Sheffield Teaching Hospital. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care or the funders.

References

1 Team WHOD of MH and SA and WHO and WHOD of MH and SAMH and WHOMHE an. *Mental health atlas 2005*. World Health Organisation 2005.

2 Studenski S, Perera S, Patel K, *et al.* Gait Speed and Survival in Older Adults. *JAMA* 2011;**305**:50–8. doi:10.1001/jama.2010.1923

3 Perera S, Patel K V, Rosano C, *et al.* Gait Speed Predicts Incident Disability: A Pooled Analysis. *Journals Gerontol Ser A* 2016;**71**:63–71. doi:10.1093/gerona/glv126

4 Rochester L, Mazzà C, Mueller A, *et al.* A Roadmap to Inform Development, Validation and Approval of Digital Mobility Outcomes: The Mobilise-D Approach. *Digit Biomarkers* 2020;**4**:13–27. doi:10.1159/000512513

5 Del Din S, Godfrey A, Mazzà C, *et al.* Free-living monitoring of Parkinson's disease: Lessons from the field. *Mov Disord* 2016;**31**:1293–313. doi:10.1002/mds.26718

6 Bonci T, Keogh A, Din S Del, *et al.* An objective methodology for the selection of a device for continuous mobility assessment. *Sensors (Switzerland)* 2020;**20**:1–16. doi:10.3390/s20226509

7 Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. *Ind Manag Data Syst* 2015;**115**:1704–23. doi:10.1108/IMDS-03-2015-0087

8 Jia Y, Wang W, Wen D, *et al.* Perceived user preferences and usability evaluation of mainstream wearable devices for health monitoring. *PeerJ* 2018;**2018**:1–17. doi:10.7717/peerj.5350

9 Puri A, Kim B, Nguyen O, *et al.* User acceptance of wrist-worn activity trackers among community-dwelling older adults: Mixed method study. *JMIR mHealth uHealth* 2017;**5**. doi:10.2196/mhealth.8211

10 Viceconti M, Penna SH, Dartee W, *et al.* Toward a regulatory qualification of real-world mobility performance biomarkers in parkinson's patients using digital mobility outcomes. *Sensors (Switzerland)* 2020;**20**:1–13. doi:10.3390/s20205920

11 Mobilise-D website. https://www.mobilise-d.eu/ (accessed 18 Feb 2021).

12 Bartlett JW, Frost C. Reliability, repeatability and reproducibility: Analysis of measurement errors in continuous variables. *Ultrasound Obstet Gynecol* 2008;**31**:466–75. doi:10.1002/uog.5256

13 Letter of support for Mobilise-D digital mobility outcomes as monitoring biomarkers. https://www.ema.europa.eu/en/documents/other/letter-support-mobilise-d-digital-mobility-outcomes-monitoring-biomarkers_en.pdf (accessed 18 Feb 2021).

14 Systems M, Development S, Devices E. *IEEE Standard for Sensor Performance Parameter Definitions*. 2014.

15 El-Sheimy N, Hou H, Niu X. Analysis and modeling of inertial sensors using allan variance. *IEEE Trans Instrum Meas* 2008;**57**:140–9. doi:10.1109/TIM.2007.908635

16 Hussen AA, Jleta IN. Low-Cost Inertial Sensors Modeling Using Allan Variance. *Int Sch Sci Res Innov* 2015;**9**:1069–74.

17 IEEE Aerospace and Electronic Systems Society (institution). *IEEE Standard Specification Format Guide and Test Procedure for Single-Axis Laser Gyros*. 1995.

18 Ferraris F, Gorini I, Grimaldi U, *et al.* Calibration of three-axial rate gyros without angular velocity standards. *"Sensors Actuators, A Phys* 1994;**42**:446–9. doi:10.1016/0924-4247(94)80031-6

19 Stančin S, Tomažič S. Time- and computation-efficient calibration of MEMS 3D accelerometers and gyroscopes. *Sensors (Switzerland)* 2014;**14**:14885–915. doi:10.3390/s140814885

20 Kluge F, Del Din S, Cereatti A, *et al.* Consensus based framework for digital mobility monitoring. *medRxiv*. 2021;**1**;12 .doi:https://doi.org/10.1101/2020.12.18.20248404

21 COSMIN. https://www.cosmin.nl/ (accessed 18 Feb 2021).

22 Nasreddine ZS, Phillips NA, Bédirian V, *et al.* The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc* 2005;**53**:695–9. doi:https://doi.org/10.1111/j.1532-5415.2005.53221.x

23 Postuma RB, Berg D, Stern M, *et al.* MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2015;**30**:1591–601. doi:10.1002/mds.26424

24 Jette AM, Haley SM, Coster WJ, *et al.* Late Life Function and Disability Instrument: I. Development and Evaluation of the Disability Component. *Journals Gerontol Ser A* 2002;**57**:M209–16. doi:10.1093/gerona/57.4.M209

25 Haley SM, Jette AM, Coster WJ, *et al.* Late Life Function and Disability Instrument: II. Development and Evaluation of the Function Component. *Journals Gerontol Ser A* 2002;**57**:M217–22. doi:10.1093/gerona/57.4.M217

26 Goetz CG, Tilley BC, Shaftman SR, *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord* 2008;**23**:2129–70. doi:https://doi.org/10.1002/mds.22340

27 Kurtzke JF. Rating neurologic impairment in multiple sclerosis. *Neurology* 1983;**33**:1444 LP – 1444. doi:10.1212/WNL.33.11.1444

28 Gupta N, Pinto LM, Morogan A, *et al.* The COPD assessment test: a systematic review. *Eur Respir J* 2014;**44**:873 LP – 884. doi:10.1183/09031936.00025214

29 Green CP, Porter CB, Bresnahan DR, *et al.* Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. *J Am Coll Cardiol* 2000;**35**:1245–55. doi:10.1016/S0735-1097(00)00531-3

30 Chiari L, Della Croce U, Leardini A, *et al.* Human movement analysis using stereophotogrammetry: Part 2: Instrumental errors. *Gait & posture* 2005; **21;**197-211. doi:10.1016/j.gaitpost.2004.04.004

31 Di Marco R, Rossi S, Castelli E, *et al.* Effects of the calibration procedure on the metrological performances of stereophotogrammetric systems for human movement analysis. *Meas J Int Meas Confed* 2017;**101**:265–71. doi:10.1016/j.measurement.2016.01.008

32 Trojaniello D, Cereatti A, Croce U Della. Gait direction of progression changes using shank worn MIMUs. In: *GBN Conference, Pavia*. 2014.

33 Salis F, Bertuletti S, Caruso M, *et al.* Multi-sensor integration and data fusion for enhancing gait assessment in and out of the laboratory. *Gait Posture* 2019;**74**:34. doi:10.1016/j.gaitpost.2019.07.493

34 Bertuletti S, Della Croce U, Cereatti A. A wearable solution for accurate step detection based on the direct measurement of the inter-foot distance. *J Biomech* 2019;**84**:274–7. doi:https://doi.org/10.1016/j.jbiomech.2018.12.039

35 Caruso, M.; Sabatini, A.M.; Laidig, D.; Seel, T.; Knaflitz, M.; Della Croce, U.; Cereatti, A. Analysis of the Accuracy of Ten Algorithms for Orientation Estimation Using Inertial and Magnetic Sensing under Optimal Conditions: One Size Does Not Fit All. *Sensors* 2021;**21**; 2543. doi:10.3390/s21072543

36 Caruso M, Sabatini AM, Knaflitz M, *et al.* Accuracy of the Orientation Estimate Obtained Using Four Sensor Fusion Filters Applied to Recordings of Magneto-Inertial Sensors Moving at Three Rotation Rates. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019;2053–8. doi:10.1109/EMBC.2019.8857655

37 Skog I, Nilsson J, Händel P. Evaluation of zero-velocity detectors for foot-mounted inertial navigation systems. In: *2010 International Conference on Indoor Positioning and Indoor Navigation*. 2010;1–6. doi:10.1109/IPIN.2010.5646936

38 Peruzzi A, Della Croce U, Cereatti A. Estimation of stride length in level walking using an inertial measurement unit attached to the foot: A validation of the zero velocity assumption during stance. *J Biomech* 2011;**44**:1991–4. doi:10.1016/j.jbiomech.2011.04.035

39 Trojaniello D, Cereatti A, Pelosin E, *et al.* Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors : application to elderly , hemiparetic , parkinsonian and choreic gait. 2014;**44**:152.

40 Köse A, Cereatti A, Della Croce U. Bilateral step length estimation using a single inertial measurement unit attached to the pelvis. *J Neuroeng Rehabil* 2012;**9**. doi:10.1186/1743-0003-9-9

41 Trojaniello D, Cereatti A, Ravaschio A, *et al.* Assessment of gait direction changes during straight-ahead walking in healthy elderly and Huntington Disease patients using a shank worn MIMU. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2014. 2508–11. doi:10.1109/EMBC.2014.6944132

42 Betoli M, Cereatti A, Trojaniello D, *et al*. Estimation of spatio-temporal parameters of gait from magneto-inertial measurement units: multicenter validation among Parkinson, mildly cognitively impaired and healthy older adults. *Biomedical engineering online.* 2018 Dec;**17**(1):1-4. doi:10.1186/s12938-018-0488-2

43 Rossaningo R, Caruso M, Salis F, *et al*. An Optimal Procedure for Stride Length Estimation Using Foot-Mounted Magneto-Inertial Measurement Units. *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)* 2021 Jun 23 (pp. 1-6). IEEE., doi: 10.1109/MeMeA52024.2021.9478604.

44 Salis F, Bertuletti S, Caruso M, *et al.* A novel multi-sensor system for gait assessment in real-world conditions: preliminary results. In: *Proceedings of GNB 2020*. 2020.

45 Zijlstra W, Hof AL. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture* 2003;**18**:1–10. doi:https://doi.org/10.1016/S0966-6362(02)00190-X

46 Ngueleu AM, Blanchette AK, Bouyer L, *et al.* Design and accuracy of an instrumented insole using pressure sensors for step count. *Sensors (Switzerland)* 2019;**19**:12–4. doi:10.3390/s19050984

47 Nightingale CJ, Mitchell SN, Butterfield SA. Validation of the Timed Up and Go Test for Assessing Balance Variables in Adults Aged 65 and Older. *J Aging Phys Act*;**27**:230–3. doi:10.1123/japa.2018-0049 10.1123/japa.2018-0049

48 Deathe AB, Miller WC. The L Test of Functional Mobility: Measurement Properties of a

Modified Version of the Timed “Up & Go” Test Designed for People With Lower-Limb Amputations. *Phys Ther* 2005;**85**:626–35. doi:10.1093/ptj/85.7.626

49 Pham MH, Elshehabi M, Haertner L, *et al.* Validation of a Step Detection Algorithm during Straight Walking and Turning in Patients with Parkinson’s Disease and Older Adults Using an Inertial Measurement Unit at the Lower Back . Front. Neurol. . 2017;**8**:457.https://www.frontiersin.org/article/10.3389/fneur.2017.00457

50 Ciravegna F, Gao J, Ireson N, *et al.* Active 10: Brisk Walking to Support Regular Physical Activity. In: *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. New York, NY, USA: : Association for Computing Machinery 2019. 11–20. doi:10.1145/3329189.3329208

51 Open Street Map. https://www.openstreetmap.org (accessed 18 Feb 2021).

52 Ordnance Survey Terrain 5. https://www.ordnancesurvey.co.uk/business-government/products/terrain-5 (accessed 18 Feb 2021).

53 U.S. Releases Enhanced Shuttle Land Elevation Data. https://www2.jpl.nasa.gov/srtm/ (accessed 18 Feb 2021).

54 Rabinovich RA, Louvaris Z, Raste Y, *et al.* Validity of physical activity monitors during daily life in patients with COPD. *Eur Respir J* 2013;**42**:1205–15. doi:10.1183/09031936.00134312

55 Knight JF, Baber C, Schwirtz A, *et al.* The comfort assessment of wearable computers. *Proc - Int Symp Wearable Comput ISWC* 2002;**2002-Janua**:65–72. doi:10.1109/ISWC.2002.1167220

56 Brooke J. SUS: a “quick and dirty’usability". *Usability Eval Ind* 1196;**Jun 11**:189.

57 Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *Int J Human–Computer Interact* 1995;**7**:57–78. doi:10.1080/10447319509526110

58 Hiden H, Woodman S, Watson P, *et al.* Developing cloud applications using the e-Science Central platform. *Philos Trans R Soc A Math Phys Eng Sci* 2013;**371**. doi:10.1098/rsta.2012.0085

59 Data Integrity and Compliance With CGMP. Guidance for Industry. https://www.fda.gov/files/drugs/published/Data-Integrity-and-Compliance-With-Current-Good-Manufacturing-Practice-Guidance-for-Industry.pdf (accessed 18 Feb 2021).

60 Tietsch M, Muaremi A, Clay I, *et al*. Robust Step Detection from Different Waist-Worn Sensor Positions: Implications for Clinical Studies. *Digital biomarkers* 2020;4(1):50-8. doi:10.1159/000511611

61 Kang H. The prevention and handling of the missing data. *Korean journal of anesthesiology*.

2013 May;64(5):402. doi:10.4097/kjae.2013.64.5.402

62 Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101. doi:10.1191/1478088706qp063oa

Figure legends

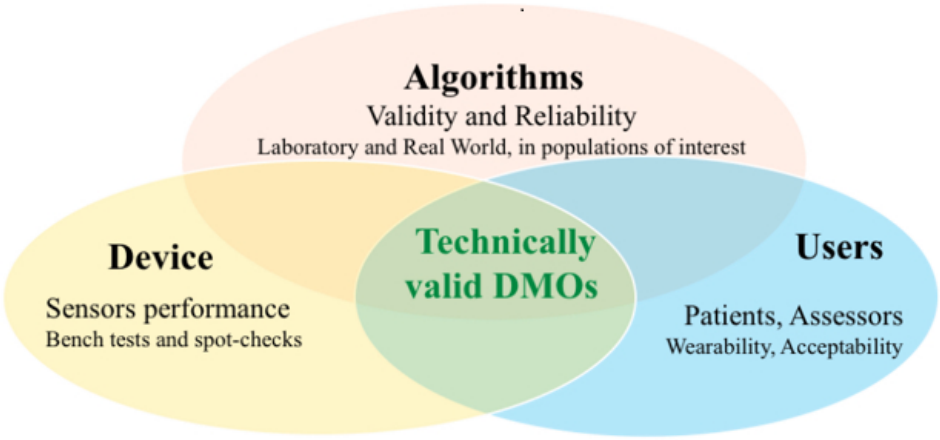
Figure 1 – Concurrent domains to be assessed as part of a technical validation of digital mobility outcomes (DMOs) obtained from a wearable device.

Figure 2 – Testing configurations used during (a) short static (plexiglass cube and device) and (b) dynamic (turntable and device) acquisitions. Marks (in red) are applied on both on the turntable and on the base to identify start/end, which have to align for the dynamic tests.

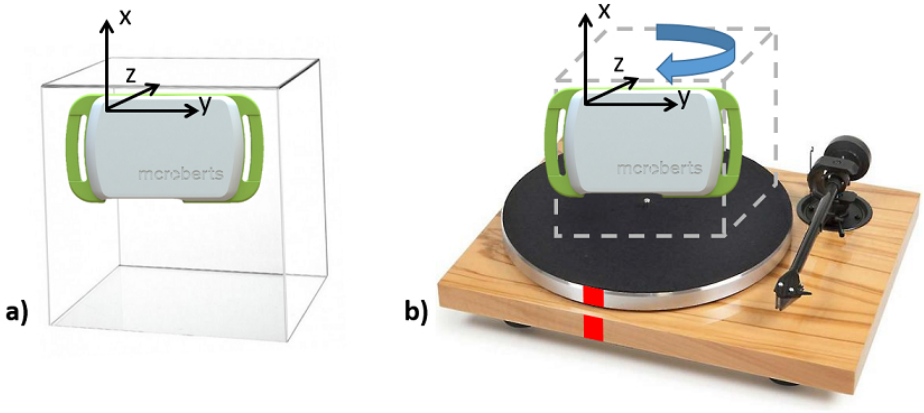
Figure 3 – Illustration of the adopted marker set configuration. Markers were located on the right (RHEEL) and left (LHEEL) heels, toes (RTOE, LTOE) and on the INDIP units located on the right and left foot (RINDIP, LINDIP). Two additional reference markers were asymmetrically attached to the side of the foot to favour automatic recognition (RREF. LREF). Four additional markers were located on the DynaPort MM+ sensor (DYNAY, DYNAO, DYNAX, DYNAREF).

Figure 4 –Different components of the INDIP system. The figure on the left shows the pressure insoles and the connectors that link them to the distance sensors and the inertial modules. The picture on the right shows how the same system is then attached to the participant's foot and leg.

Figure 5 – Diagrams of the selected tasks: (A) Straight Walking Test, (B) Timed Up and Go, (C) L-Test, (D) Surface Test, (E) Hallway Test, F) Schematic of the Daily living activities, G) Description of the eight tasks performed during the Daily Living activities.



Concurrent domains to be assessed as part of a technical validation of digital mobility outcomes (DMOs) obtained from a wearable device.



Testing configurations used during (a) short static (plexiglass cube and device) and (b) dynamic (turntable and device) acquisitions. Marks (in red) are applied on both on the turntable and on the base to identify start/end, which have to align for the dynamic tests.

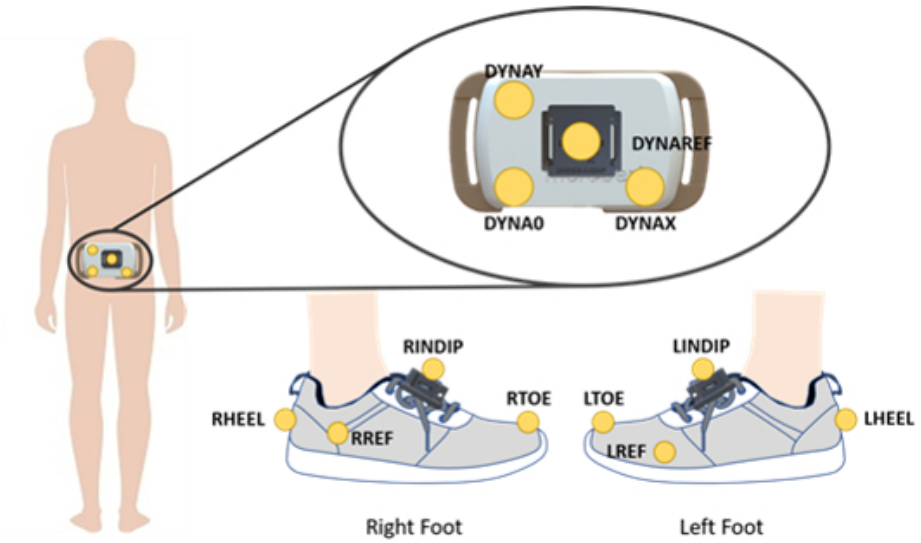
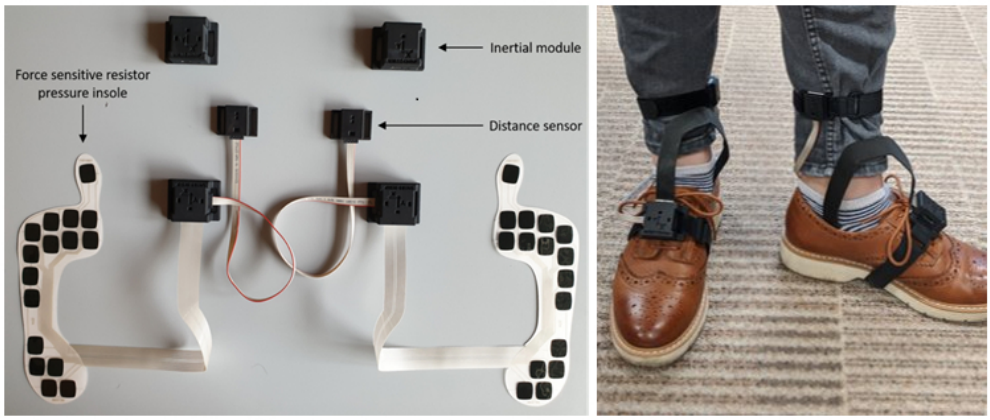
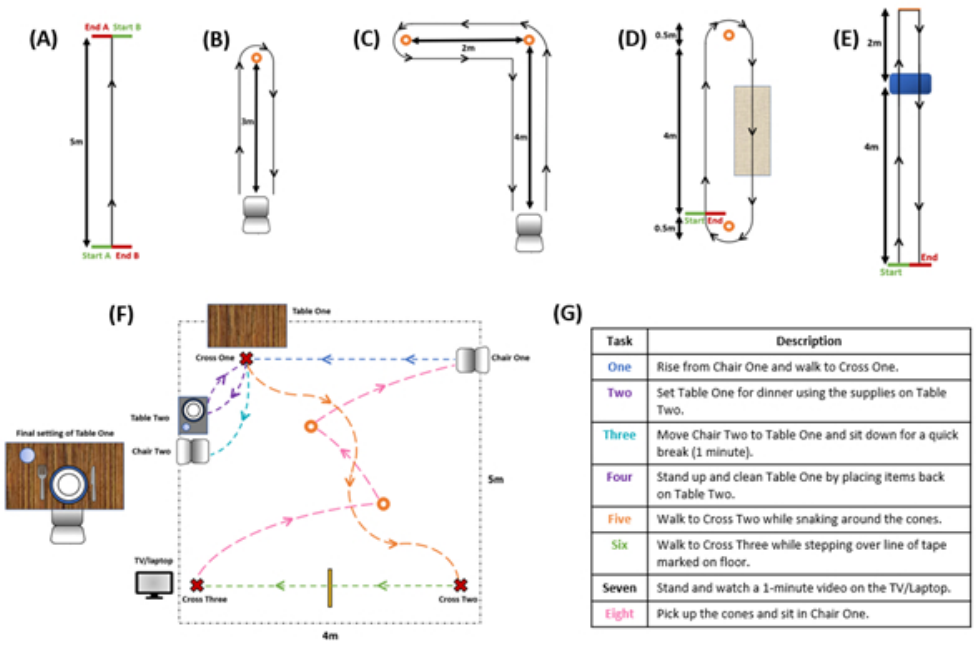


Illustration of the adopted marker set configuration. Markers were located on the right (RHEEL) and left (LHEEL) heels, toes (RTOE, LTOE) and on the INDIP units located on the right and left foot (RINDIP, LINDIP). Two additional reference markers were asymmetrically attached to the side of the foot to favour automatic recognition (RREF. LREF). Four additional markers were located on the DynaPort MM+ sensor (DYNAY, DYNAO, DYNAX, DYNAREF).



Different components of the INDIP system. The figure on the left shows the pressure insoles and the connectors that link them to the distance sensors and the inertial modules. The picture on the right shows how the same system is then attached to the participant's foot and leg.



(G)

Task	Description
One	Rise from Chair One and walk to Cross One.
Two	Set Table One for dinner using the supplies on Table Two.
Three	Move Chair Two to Table One and sit down for a quick break (1 minute).
Four	Stand up and clean Table One by placing items back on Table Two.
Five	Walk to Cross Two while snaking around the cones.
Six	Walk to Cross Three while stepping over line of tape marked on floor.
Seven	Stand and watch a 1-minute video on the TV/Laptop.
Eight	Pick up the cones and sit in Chair One.

Diagrams of the selected tasks: (A) Straight Walking Test, (B) Timed Up and Go, (C) L-Test, (D) Surface Test, (E) Hallway Test, F) Schematic of the Daily living activities, G) Description of the eight tasks performed during the Daily Living activities.

159x102mm (96 x 96 DPI)

Mobilise-D Technical Validation Study

Participant interview topic guide

Interview guide

Note: Questions in this topic guide are included to answer the above aims. If the participant is open and talks freely, they may answer some of the questions without being asked. Therefore, depending on the person, not all of these questions need to be asked. If they begin to talk about topics that may be interesting or relevant to the above aims, please feel free to continue to explore these, even if there is no specific question linked to it. In contrast, if participants are not very open, some potential prompts have been included with the questions below. These prompts are there as an optional guide and do not need to be used.

Dynaport questions	
Aim: Ensure that the McRoberts Dynaport device is comfortable and acceptable to participants	
Main questions	
1	Can you describe your experience of using the Dynaport sensor in the last week?
2	Can you tell me what you liked about the device? Disliked? *Prompts if needed* - Size/weight - Attachment to body - Ease of use - Comfort
3	How did the device make you feel? / Can you describe what it felt like to wear the device? *Prompts if needed* - In social environments/at home - Interaction with daily activities - Emotions associated with being monitored
4	What you change about the device if you could?
Follow-up questions **"**If you don't mind, I'd like to ask you some specific details about the device."	
1	How did you find the process of putting it on and taking it off?
2	How did the device influence your daily activities? *Prompts if needed* - How were they impacted?

- How did this make them feel?

3 Can you tell me about any difficulties that you had with the device?

4 How did you feel about wearing the device for a week?

Prompts if needed

- How would they feel if it was longer?
- Any concerns for the week?

Closing question

1 Is there anything else you would like me to know about the Dynaport?

The use of wearable devices in healthcare questions

Aim: Explore the acceptability of wearable devices, for the purposes of healthcare monitoring, to participants in general.

Main questions

1 Can you tell me about your experience of your health condition?

Note: Condition specific symptoms are listed at the end of this document.

2 Can you tell me about your experience of the care you've received for your condition?

Prompts if needed

- How do they feel about it?

3 Can you tell me about what sort of technology you currently use in your everyday life?

Prompts if needed

- How do you feel about using technology?
- What would make you use technology more? Less?
- Emotions associated with being monitored

4 What are your opinions on the use of technology in healthcare?

Prompts if needed

- What do you think it can be used for?

5 How would you feel about using technology to generate health information about yourself? (e.g. condition related smartphone app, self-reported outcomes platform, fitness tracker etc.),

Prompts if needed

- Why?
- What would make you use it?
- What would stop you from using it?

- What would need to change for you to use it?

Follow-up questions	
1	How do you feel about capturing health information in your daily life, using a wearable remote monitoring device?
2	How do you think digital technology used in your daily life would influence how you manage your condition? *Prompts if needed* - How would it impact their relationship with their health care provider? - Integration into activities of daily living
3	How would you feel about sharing this data with your health care provider? What about researchers? *Prompts if needed* - Usefulness - Impact of this

Closing question	
1	Is there anything else you would like me to know about using technology in healthcare?

Researcher feedback

Questionnaires

The System Usability Scale

To be completed by researchers who have collected data using the McRoberts device. When answering these questions please consider how you personally have found the McRoberts device, not how the participants have found it.

	1 strongly disagree	2	3	4	5 strongly agree
I think that I would like to use this system frequently.					
I found the system unnecessarily complex.					
I thought the system was easy to use.					

I think that I would need the support of a technical person to be able to use this system.					
I found the various functions in this system were well integrated.					
I thought there was too much inconsistency in this system.					
I would imagine that most people would learn to use this system very quickly.					
I found the system very cumbersome to use.					
I felt very confident using the system.					
I needed to learn a lot of things before I could get going with this system.					
I think that I would like to use this system frequently.					

The IBM Computer System Usability Questionnaire

To be completed by researchers who have collected data using the McRoberts device. When answering these questions please consider how you personally have found the McRoberts device, not how the participants have found it.

	1 strongly disagree	2	3	4	5	6	7 Strongly agree
Overall, I am satisfied with how easy it is to use this system.							
It is simple to use this system.							
I can effectively complete my work using this system.							

I am able to complete my work quickly using this system.							
I am able to efficiently complete my work using this system.							
I feel comfortable using this system.							
It was easy to learn to use this system.							
I believe I became productive quickly using this system.							
The system gives error messages that clearly tell me how to fix problems.							
Whenever I make a mistake using the system, I recover easily and quickly.							
The information (such as on-line help, on-screen messages and other documentation) provided with this system is clear.							
It is easy to find the information I need.							
The information provided with the system is easy to understand.							
The information is effective in helping me complete my work.							
The organization of information on the system screens is clear.							
The interface of this system is pleasant.							
I like using the interface of this system.							
This system has all the functions and capabilities I expect it to have.							

Intervention specific questionnaire

This questionnaire aims to evaluate your experiences and opinions on the processes of the Mobilise-D technical validation study. In particular, we are looking to focus on the components of the study which will also take place within the clinical validation trial. Therefore, when asking about certain aspects of the trial, we may be specific in relation to which device or measurement tool we want

feedback on. Please read each question carefully to ensure that your answers relate to the specific component under investigation.

The purpose of this questionnaire is to determine your opinions on i) the training you received, ii) the supporting materials that were provided to you, iii) the feasibility of conducting participant recruitment. Please answer each question honestly. All responses will be de-identified and treated anonymously.

Training

When considering training, we want you to consider the training that you received to carry out the following:

- Recruit participants
- Complete the human factors assessment
- Use the McRoberts device
- Use the ERT platform
- Use the EScience platform

Many of these training components would have been delivered separately. We want you to consider your experience on these training sessions overall, and then answer questions for each specific component independently. There is space to write comments at the end of this section if you wish to add more information or clarify any aspect further.

Overall training	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
The Mobilise-D training programme was enjoyable						
The Mobilise-D training programme was useful						
The Mobilise-D training programme successfully prepared me to collect data within the technical validation trial						

The Mobilise-D training programme successfully prepared me to recruit participants within the technical validation trial						
Overall training content	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
The training was sufficiently interactive						
The training was interesting						
The training was easy to understand						
The training provided a clear outline of what was expected from me						
The training provided a clear rationale of recruitment processes						
Recruitment training	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
I felt confident in my recruitment role following training						
I felt competent to complete recruitment following training						
Following training, I had no questions regarding recruitment						
Following training I was confident in who I should						

contact if I had a question regarding recruitment						
Data collection training	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
I felt confident in my data collection role following training						
I felt competent to complete data collection following training						
Following training, I had no questions regarding data collection						
Following training I was confident in who I should contact if I had a question regarding data collection						
McRoberts training	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
I felt confident in using McRoberts following training						
I felt competent to complete data collection with McRoberts following training						
Following training, I had no questions regarding the McRoberts device						
Following training I was confident in who I should contact if I had a question regarding McRoberts						

ERT platform training	Strongly disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly agree 5	Not applicable
I felt confident in the use of the ERT platform following training						
I felt competent to complete data collection using ERT following training						
Following training, I had no questions regarding the ERT platform						
Following training I was confident in who I should contact if I had a question regarding the ERT platform						

Please provide us with any further feedback you have regarding the training that you undertook as part of the Mobilise-D project:

Project materials feedback

As part of the Mobilise-D project, you received a number of materials to help support you in the recruitment and data collection processes (i.e. manuals, etc). Please provide us with your feedback on these items

Recruitment materials	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
The recruitment materials were easy to understand						
I used the recruitment materials a lot to help me recruit participants						
As recruitment progressed I used the materials less						
The materials answered all my recruitment questions						
The materials made the recruitment process easier						
When I used the materials I felt more confident in the recruitment process						
Data collection materials	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
The data collection materials were easy to understand						
I used the data collection materials a lot to help me recruit participants						
As data collection progressed I used the materials less						
The materials answered all my data collection questions						

The materials made the data collection process easier
When I used the materials I felt more confident in the data collection process

Please provide us with any further feedback you have regarding the materials that you received as part of the Mobilise-D project:

_____

_____

_____

_____

_____

_____

_____

Feasibility of the trial procedures

Recruitment

How many participants did your site recruit?

How many participants was your site due to recruit?

What barriers to participant recruitment did you encounter?

What helped you to recruit people?

What additional materials do you believe would help support recruitment in the future?

What additional support do you believe would help support recruitment in the future?

Data collection

How many participants failed to complete full data collection procedures?

Why?

What additional materials do you believe would help support data collection in the future?

What additional support do you believe would help support data collection in the future?

Do you have any other thoughts regarding the recruitment and data collection procedures of the Mobilise-D technical validation trial?

Researcher feedback

Interview guide

Start with a general question about how they found the experience of data collection within the Mobilise-D validation study

The Dynaport device

- Tell me about how you found using the Dynaport sensor during the trial?
- What were your first impressions of the device?
- Can you tell me what you liked about the device? Disliked?
- Can you tell me about any difficulties that you had with the device or its platform during the trial?
    - Difficulties they encountered themselves
        - Set up
        - Ease of explanation for participants
    - Difficulties that were reported to them
        - Did participant get in contact during the week?
        - Did any devices come back damaged?
- How would you compare this device to other wearables that you have used before?

The ERT device

- Tell me about how you found using the ERT platform during the trial?
- What were your first impressions of the device?
- Did any of your opinions change as the trial progressed?
- Can you tell me what you liked about the device? Disliked?
- Can you tell me about any difficulties that you had with the device during the trial?
- How did the device make you feel?

The EScience platform

- Tell me about how you found using the EScience platform during the trial?

- What were your first impressions of the platform?
- Did any of your opinions change as the trial progressed?
- Can you tell me what you liked about the platform? Disliked?
- Can you tell me about any difficulties that you had with the device during the trial?

The training and materials used

- Can you tell me what your opinions are of the training you received before starting data collection?

  Specifically related to: Dynaport, EScience, ERT and recruitment procedures
  - Duration
  - Content
  - Facilitators
  - Materials provided
- What did you expect to get out of the training?
- What was useful about the training? Not useful?
- Describe how you found the process of data collection?
- How did this compare to what you expected following training?
- What would you change about the training if you could? And the materials provided?

Recruitment procedures

Can you tell me about your experiences recruiting participants for the Mobilise-D trial
What needs to change to make recruitment easier?

BMJ Open

A MULTI-CENTRIC OBSERVATIONAL STUDY PROTOCOL FOR THE TECHNICAL VALIDATION OF REAL-WORLD MONITORING OF GAIT

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-050785.R2
Article Type:	Protocol
Date Submitted by the Author:	20-Oct-2021

Complete List of Authors:	Mazzà, Claudia; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Alcock, Lisa; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Aminian, Kamiar; Ecole Polytechnique Federale de Lausanne, Laboratory of Movement Analysis and Measurement Becker, Clemens; Robert Bosch Gesellschaft für Medizinische Forschung Bertuletti, Stefano; University of Sassari, Department of Biomedical Sciences Bonci, Tecla; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Brown, Philip; Newcastle Upon Tyne Hospitals NHS Foundation Trust Brozgol, Marina; Tel Aviv Sourasky Medical Center, Center for the Study of Movement, Cognition and Mobility, Neurological Institute Buckley, Ellen; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Carsin, Anne-Elie; Instituto de Salud Global Barcelona Caruso, Marco; Politecnico di Torino Dipartimento di Elettronica e Telecomunicazioni; Politecnico di Torino, PolitoBIOMed Lab – Biomedical Engineering Lab Caulfield, Brian; University College Dublin, Insight Centre for Data Analytics, O'Brien Science Centre; University College Dublin, UCD School of Public Health, Physiotherapy and Sports Science Cereatti, Andrea; Politecnico di Torino Dipartimento di Elettronica e Telecomunicazioni Chiari, Lorenzo; University of Bologna, Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi»; University of Bologna, Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV) Chynkiamis, Nikolaos; Northumbria University Newcastle, Department of Sport, Exercise and Rehabilitation Ciravegna, Fabio; The University of Sheffield, Department of Computer Science; The University of Sheffield, INSIGNEO Institute for in silico Medicine Del Din, Silvia; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences

	Eskofier, Björn; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering Evers, Jordi; McRoberts BV Garcia Aymerich, Judith; Instituto de Salud Global Barcelona Gazit, Eran; Tel Aviv Sourasky Medical Center, Center for the Study of Movement, Cognition and Mobility, Neurological Institute Hansen, Clint; University Medical Center Schleswig-Holstein Campus Kiel, Department of Neurology Hausdorff, Jeffrey; Tel Aviv Sourasky Medical Center, Center for the Study of Movement, Cognition and Mobility, Neurological Institute; Tel Aviv University, Department of Physical Therapy, Sackler Faculty of Medicine & Sagol School of Neuroscience Helbostad, Jorunn; Norwegian University of Science and Technology, Department of Neuromedicine and Movement Science Hiden, Hugo; University of Newcastle upon Tyne, School of Computing Hume, Emily; Northumbria University Newcastle, Department of Sport, Exercise and Rehabilitation Paraschiv-Ionescu, Anisoara; Ecole Polytechnique Federale de Lausanne, Laboratory of Movement Analysis and Measurement Ireson, Neil; The University of Sheffield, Department of Computer Science; The University of Sheffield, INSIGNEO Institute for in silico Medicine Keogh, Alison; University College Dublin, Insight Centre for Data Analytics, O'Brien Science Centre; University College Dublin, UCD School of Public Health, Physiotherapy and Sports Science Kirk, Cameron; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Kluge, Felix; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering Koch, Sarah; Instituto de Salud Global Barcelona Küderle, Arne; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering Lanfranchi, Vitaveska; The University of Sheffield, Department of Computer Science; The University of Sheffield, INSIGNEO Institute for in silico Medicine Maetzler, Walter; University Medical Center Schleswig-Holstein Campus Kiel, Department of Neurology Micó-Amigo, M. Encarna; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Mueller, Arne; Novartis Pharma AG, Novartis Institutes of Biomedical Research Neatrour, Isabel; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences Niessen, Martjin; McRoberts BV Palmerini, Luca; University of Bologna, Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi»; University of Bologna, Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV) Pluimgraaff, Lucas; McRoberts BV Reggi, Luca; University of Bologna, Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV) Salis, Francesca; University of Sassari, Department of Biomedical Sciences Schwickert, Lars; Robert Bosch Gesellschaft für Medizinische Forschung Scott, Kirsty; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Sharrack, Basil; Sheffield Teaching Hospitals NHS Foundation Trust,

	Department of Neuroscience and Sheffield NIHR Translational Neuroscience BRC Sillen, Henrik; AstraZeneca Sweden, Digital Health R&D Singleton, David; University College Dublin, Insight Centre for Data Analytics, O'Brien Science Centre; University College Dublin, UCD School of Public Health, Physiotherapy and Sports Science Soltani, Abolfazi; Ecole Polytechnique Federale de Lausanne, Laboratory of Movement Analysis and Measurement Taraldsen, Kristin; Norwegian University of Science and Technology, Department of Neuromedicine and Movement Science Ullrich, Martin; Friedrich-Alexander-Universität Erlangen-Nürnberg, Machine Learning and Data Analytics Lab, Department of Computer Science Van Gelder, Linda; The University of Sheffield, INSIGNEO Institute for in silico Medicine; The University of Sheffield, Department of Mechanical Engineering Vereijken, Beatrix; Norwegian University of Science and Technology, Department of Neuromedicine and Movement Science Vogiatzis, Ioannis ; Northumbria University Newcastle, Department of Sport, Exercise and Rehabilitation Warmerdam, Elke; University Medical Center Schleswig-Holstein Campus Kiel, Department of Neurology Yarnall, Alison; Newcastle University Faculty of Medical Sciences, Translational and Clinical Research Institute; Newcastle Upon Tyne Hospitals NHS Foundation Trust Rochester, Lynn; University of Newcastle upon Tyne, Translational and Clinical Research Institute, Faculty of Medical Sciences; Newcastle Upon Tyne Hospitals NHS Foundation Trust
Primary Subject Heading:	Diagnostics
Secondary Subject Heading:	Research methods, Neurology, Respiratory medicine
Keywords:	Multiple sclerosis < NEUROLOGY, Parkinson-s disease < NEUROLOGY, Hip < ORTHOPAEDIC & TRAUMA SURGERY, Chronic airways disease < THORACIC MEDICINE, Heart failure < CARDIOLOGY

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

A MULTI-CENTRIC OBSERVATIONAL STUDY FOR THE TECHNICAL VALIDATION OF REAL-WORLD MONITORING OF GAIT

Mazzà Claudia[1,2], Alcock Lisa[3], Aminian Kamiar[4], Becker Clemens[5], Bertuletti Stefano[6], Bonci Tecla[1,2], Brown Philip[7], Brozgol Marina[8], Buckley Ellen[1,2], Carsin Anne-Elie[9], Caruso Marco[10,11], Caulfield Brian[12,13], Cereatti Andrea[10], Chiari Lorenzo[14,15], Chynkiamis Nikolaos[16], Ciravegna Fabio[1,17], Del Din Silvia[3], Eskofier Björn[18], Evers Jordi[19], Garcia-Aymerich, Judith[9], Gazit Eran[8], Hansen Clint[20], Hausdorff Jeffrey[8,21], Helbostad Jorunn[22], Hiden Hugo[23], Hume Emily[16], Paraschiv-Ionescu Anisoara[4], Ireson Neil[1,17], Keogh Alison[12,13], Kirk Cameron[3], Kluge Felix[18], Koch Sarah[9], Küderle Arne[18], Lanfranchi Vitaveska[1,17], Maetzler Walter[20], Micó-Amigo M. Encarna[3], Mueller Arne[24], Neatrour Isabel[3], Niessen Martjin[19], Palmerini Luca[14,15], Pluimgraaff Lucas[19], Reggi Luca[15], Salis Francesca[6], Schwickert Lars[5], Scott Kirsty[1,2], Sharrack Basil[25], Sillen Henrik[26], Singleton David[12,13], Soltani Abolfazi[4], Taraldsen Kristin[22], Ullrich Martin[18], Van Gelder Linda[1,2], Vereijken Beatrix[22], Vogiatzis, Ioannis[16], Warmerdam Elke[20], Yarnall Alison[3,7], Rochester Lynn[3,7]

Corresponding author:
Prof. Claudia Mazzà
c.mazza@sheffield.ac.uk
Pam Liversidge Building - Room F14
Mappin Street, Sheffield, S1 3JD, UK
Phone: +44 (0)114 2226073

1. INSIGNEO Institute for in silico Medicine, The University of Sheffield, Sheffield, UK
2. Department of Mechanical Engineering, The University of Sheffield, Sheffield, UK
3. Translational and Clinical Research Institute, Faculty of Medical Sciences, University of Newcastle upon Tyne, Newcastle upon Tyne, UK
4. Laboratory of Movement Analysis and Measurement, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland
5. Robert Bosch Gesellschaft für Medizinische Forschung, Stuttgart, Germany
6. Department of Biomedical Sciences, University of Sassari, Sassari, Italy
7. Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK
8. Center for the Study of Movement, Cognition and Mobility, Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel
9. Instituto de Salud Global Barcelona, Barcelona, Spain
10. Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Turin, Italy
11. PolitoBIOMed Lab – Biomedical Engineering Lab, Politecnico di Torino, Turin, Italy
12. Insight Centre for Data Analytics, O'Brien Science Centre, University College Dublin, Dublin, Ireland
13. UCD School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland
14. Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi», University of Bologna, Bologna, Italy
15. Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV), University of Bologna, Bologna, Italy
16. Department of Sport, Exercise and Rehabilitation, Northumbria University Newcastle, Newcastle upon Tyne, UK
17. Department of Computer Science, The University of Sheffield, Sheffield, UK

18. Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
19. McRoberts BV, The Hague, Netherlands
20. Department of Neurology, University Medical Center Schleswig-Holstein Campus Kiel, Kiel, Germany
21. Department of Physical Therapy, Sackler Faculty of Medicine & Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel
22. Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway
23. School of Computing, University of Newcastle upon Tyne, Newcastle upon Tyne, UK
24. Novartis Institutes of Biomedical Research, Novartis Pharma AG, Basel, Switzerland
25. Department of Neuroscience and Sheffield NIHR Translational Neuroscience BRC, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
26. Digital Health R&D, AstraZeneca, Sweden

Keywords

Gait analysis; wearable sensors; activity monitoring; users evaluation; mobility outcomes

Wordcount: 5491

1 ABSTRACT

Introduction

Existing mobility endpoints based on functional performance, physical assessments and patient self-reporting are often affected by lack of sensitivity, limiting their utility in clinical practice. Wearable devices including inertial measurement units (IMUs) can overcome these limitations by quantifying digital mobility outcomes (DMOs) both during supervised structured assessments and in real-world conditions. The validity of IMU-based methods in the real-world, however, is still limited in patient populations. Rigorous validation procedures should cover the device metrological verification, the validation of the algorithms for the DMOs computation specifically for the population of interest and in daily life situations, and the users' perspective on the device.

Methods and analysis

This protocol was designed to establish the technical validity and patient acceptability of the approach used to quantify digital mobility in the real-world by Mobilise-D, an EU-funded consortium part of the Innovative Medicine Initiative, aiming at fostering regulatory approval and clinical adoption of DMOs. After defining the procedures for the metrological verification of an IMU based device, the experimental procedures for the validation of algorithms used to calculate the DMOs are presented. These include laboratory and real-world assessment in 120 participants from five groups: healthy older adults; chronic obstructive pulmonary disease, Parkinson's disease, multiple sclerosis, proximal femoral fracture, and congestive heart failure. DMOs extracted from the monitoring device will be compared to those from different reference systems, chosen according to the contexts of observation. Questionnaires and interviews will evaluate the users' perspective on the deployed technology and relevance of the mobility assessment.

Ethics and Dissemination

The study, registered at ISRCTN (12246987), has been granted ethics approval by the centre's committees (London – Bloomsbury Research Ethics committee; Helsinki Committee, Tel Aviv Sourasky Medical Centre; Medical Faculties of The University of Tübingen and of the University of Kiel). Data and algorithms will be made publicly available.

Strengths and limitations of this study

- A multi-disciplinary approach was implemented to define a protocol for the validation of tools for mobility monitoring, covering aspects related to devices, algorithms and users.
- A set of rigorous quality assurance procedures have been established to allow for the creation of a high-quality annotated dataset to foster development in the field of digital mobility monitoring.

- Mobility data will be collected in the laboratory and in the real-world for five different cohorts of slow-walkers. and subsequently will be made publicly available at the end of the study.
- The multi-stage and multi-device experimental procedures required by this validation study can be extremely challenging for both the participants and the assessors.
- For the laboratory acquisitions, the level of agreement between the gold standard and the inertial sensor devices might be affected by the limitations associated with a restricted capture volume.

2 INTRODUCTION

The ability to move is a key contributor to physical, mental and social well-being, which is in line with the World Health Organisation's (WHO's) definition of health [1]. However, the study of mobility has received relatively little attention, except for diseases characterised by specific mobility dysfunction. The increasing longevity of the world's population together with prolonged survival of many patients with long term conditions means that more people are suffering from loss of mobility, which in turn is a major contributing factor to a loss of independence [2,3]. This has a considerable and growing personal, societal and economic impact. Efforts to mitigate this loss of mobility are an increasing priority and promising interventions are now under investigation.

Existing mobility endpoints based on performance, patient self-reporting and one-off assessment are resource-intensive and lack sensitivity [4], which limits therapeutic development and clinical management. A novel approach is needed that is low cost, simple, accurate and that can be used in the real world, including the home and the community.

Poor gait, especially slow walking, is a key determinant of mobility loss. It is associated with greater mortality, morbidity, cognitive decline, dementia and fall risk [2,3]. Quantifying gait related mobility outcomes, including features such as step/stride duration and their variability, walking speed and asymmetry features is well established in supervised instrumented assessments.

Wearable devices including inertial measurement units (IMUs) that allow digital mobility outcomes (DMOs) to be described, are leading the transition from laboratory-based assessment of mobility (mobility capacity), to continuous, unsupervised monitoring of mobility in daily life conditions (mobility performance). Nonetheless, the validity of IMU-based methods to characterise real-world mobility, and gait in particular, is still limited, especially in populations suffering from pathological conditions. This is because measuring real-world gait is far from simple or straightforward. In addition, complex factors arise from multiple sources that influence outcome measures, including disease characteristics, patient specific habits, environment/context and the purpose of walking. All these factors limit the validity of existing algorithms developed to quantify targeted DMOs [5]. Additionally, validation should include simultaneous evaluation [6] of the participants perception and acceptability of the device [7] as well as aspects related to wearability and usability [8,9]. Finally, a separate assessment of the metrological performance of the sensors contained in the adopted device is required.

All these validation steps need to be taken before the DMOs and the associated technologies can be effectively used for clinical [4] and regulatory [10] purposes (see Figure 1 for summary).

[Figure 1 around here]

In this paper, we present the comprehensive, multi-stage protocol deployed in the technical validation study (TVS) that we have developed as part of the IMI2-JU-funded Mobilise-D project (Number 820820 [11]), that aims to validate a new digital method for remote monitoring of mobility. In particular, this multi-stage protocol aims to: a) verify the metrological performance of the sensors included in a IMU based monitoring device, using a procedure that could be replicated on any device; b) establish the validity and reliability [12] of the DMOs estimated by the algorithms using data from an IMU based device, taking into accounts the effects of populations (e.g., healthy adults, patients with various conditions), locomotor activities (simple straight walking versus complex walking tasks), contexts (lab based versus real-world), durations (device wearing time, DMOs hourly and daily fluctuations, etc.), and contextual confounding factors (such as location of walks, weather, use of walking aids, etc.); and, c) establish participants' and assessors' opinions on the usability and acceptability of the monitoring devices that will be deployed.

This protocol will provide the stakeholders, the first comprehensive, multi-modal solution to validate real-world mobility. Notably, the protocol described here has been accepted by the European Medical Agency [13] as part of the Mobilise-D process for the regulatory qualification of real-world mobility performance biomarkers in Parkinson's disease. [10]

3 METHODS

The monitoring device that will be used in this study to collect mobility data is the DynaPort MM+ IMU (McRoberts, Table 1). In a trade-off between usability, accuracy and ability to provide DMOs for both at step and stride level, this is attached to the lower back via an elastic waistband and Velcro strap. The validity of algorithms to accurately estimate DMOs in the real-world will be investigated in 120 participants including: healthy older adults (HA) and in five clinical cohorts (chronic obstructive pulmonary disease, COPD; Parkinson's disease, PD; multiple sclerosis, MS; proximal femoral fracture, PFF; and congestive heart failure, CHF), chosen as presenting a variety of gait and mobility features [4]. The usability and acceptability of the device from the perspective of the participants and the assessors involved in the study will be established via interviews and questionnaires.

3.1 Verification of the metrological performances of the device

Verifying the performance of a device needs a robust and comprehensive metrological characterisation of all the sensors that it embeds. This requires a series of standardised procedures (spot-checks) to be

implemented to ensure accuracy of raw data that will be used as input to the algorithms. Table 1 shows the sensing characteristics of the device used in this study, the Dynaport MM+ (dimensions: 106.6 x 58 x 11.5mm, size: 55 grams).

Table 1 Characteristics of the sensors included in the DynaPort MM+ device

Sensor	Sampling Frequency	Sensor Range	Sensor resolution
Tri-axial accelerometer	100 Hz	± 8g	1 mg (at ± 8g)
Tri-axial gyroscope	100 Hz	± 2000 dps	70 mdps (at ± 2000 dps)

According to Institute of Electrical and Electronics Engineers *Standard for Sensor Performance Parameter Definitions ((IEEE 2700-2017* [14]) a number of parameters are needed to characterise the metrological performance of the accelerometer, gyroscope, and magnetometer included in an IMU. These parameters can be computed under both static and dynamic conditions. The main noise parameters used are the first (mean value) and second (variance) order statistics, and the root Allan variance parameters of noise [15,16]. The parameters related to the first and second order statistics of noise can also be estimated by means of a short static acquisition (the minimum length of each acquisition is defined by IEEE standard for each sensor). These short static acquisitions can be performed simply using a plastic cube, where the device can be properly secured and then each of the three sensors' coordinate axes x, y and z are in turn aligned with the direction of gravity (g) as well as its opposite direction (six combinations, Figure 2). The root Allan variance parameters are instead computed over a long static acquisition [17]. Typically, the acquisitions are performed over a period of 4-8 hours [15,16]. All the static acquisitions should be carried out at a constant temperature of 25°C.

In comparison to a static acquisition, characterising the dynamic metrological performance of the sensors embedded in an IMU is less straightforward, since the metrological standards provided by IEEE describe a sequence of operations requiring an expensive and complex testing instrumentation. However, various alternatives have been proposed in the literature. The accuracy of a gyroscope can be quantified during a single-axis rotation by a known angle by computing the ideal angular velocity and comparing it to the average measured angular velocity [18,19]. This procedure should be performed using a rotation plate with a rotating speed comparable to what is encountered during human gait (~200 deg/s) (Figure 2).

[Figure 2 around here]

The tests described above will be performed on thirty-five different DynaPort MM+ devices deployed in the study. This will allow conformity with manufacturer indications to be verified, highlight the need for sensor recalibrations and provide benchmark standards for any device with equivalent sensing capacity. In turn, this will allow any device with an equivalent or superior solution to be used, in order

to facilitate broader adoption of validated algorithms and cope with a continuously changing hardware landscape.

3.2 Protocol for Validation of the algorithms

Several algorithms to detect the DMOs from a single device have been implemented according to agreed definitions [20] and based on existing literature. At this time, these algorithms are being concurrently validated using the approach described by Bonci et al. [6] using pre-existing datasets, which mostly include lab-based observations only. Following this selection process, the best performing algorithms will be assessed using the data captured with the protocol here described.

3.2.1 Ethics and Dissemination

This multi-centre study is sponsored and coordinated by The Newcastle upon Tyne Hospitals NHS Foundation Trust, UK. Participants will be recruited in five sites across Europe: Tel Aviv Sourasky Medical Center, Israel (ethics approval granted by the Helsinki Committee, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel, 0551-19TLV), Robert Bosch Foundation for Medical Research, Germany (ethics approval granted by the ethical committee of the medical faculty of The University of Tübingen, 647/2019BO2), University of Kiel, Germany (ethics approval granted by the ethical committee of the medical faculty of Kiel University, D438/18), The Newcastle upon Tyne Hospitals NHS Foundation Trust, UK and Sheffield Teaching Hospitals NHS Foundation Trust, UK (ethics approval granted by London – Bloomsbury Research Ethics committee, 19/LO/1507).

As per the study register (ISRCTN, 12246987) the data collection was originally planned to start in April 2021 and last six months. The pandemic situation meant that the study started in July 2020 and is now planned to finish in September 2021. All experimental procedures are constantly monitored and revised as needed to ensure full compliance with COVID19-related health and safety measures and for safeguarding of both the study participants and the assessors. The data collected as part of this protocol will be made publicly available together with the algorithms used to process the data.

3.2.2 Participants

A convenience sample of 120 participants will be recruited via their clinical care team or research registers to represent the five disease cohorts (COPD, PD, MS, PFF, and CHF), as well as healthy older adults (HA). Twenty participants will be recruited for each cohort. Each cohort will be recruited across multiple sites to ensure generalisability (e.g. differing cultures and contexts). Inclusion and exclusion criteria, grouped by total cohort and disease cohort, are summarised in Table 2. Given the novelty of the data, rather than on a power calculation the sample size of 120 has been initially defined according to Consensus-based Standards for the selection of health Measurement Instruments guidelines for measurement properties (COSMIN [21]). This sample size, however, will be refined after 50% of the

data collection. Given the DMOs are measured at walking bout level and not at patient level, in this analysis we will use the effective number of walking bouts observed during the 2.5 hours to perform the power calculation. We will base this analysis on a desired ICC coefficient ≥0.7, with Alpha=0.05 and Beta=0.9, and an aimed confidence interval of 0.1. Based on this review, more participants may be recruited.

All participants will give written informed consent prior to undergoing a clinic/laboratory-based session to record generic and disease-specific characteristics. This will include participant reported outcomes, assessments and medical notes review. The generic and cohort-specific clinical outcomes that will be collected are summarised in Table 3.

Table 2 Inclusion and exclusion criteria adopted for the different disease cohorts

Group	Inclusion criteria	Exclusion criteria
All groups	-able to walk 4 meters independently with or without walking aids -able to give informed consent -willingness to wear the sensor set-ups during the study -shoe size 36 EU (3 UK) or above -able to read and write in first language of the respective country -Montreal Cognitive Assessment (MoCA) >15 [22] -available for home /office visit during study period	-occurrence of any of the following 3 months prior to inclusion: myocardial infarction, hospitalisation for unstable angina, stroke, coronary artery bypass graft, percutaneous coronary intervention, implantation of a cardiac resynchronisation therapy device -current medical condition that could interfere with the patient's compliance
COPD	- ≥45 years of age -Diagnosis of COPD (post-bronchodilator forced expiratory volume in the first second (FEV_1) to forced vital capacity (FVC) ratio <0.70) -clinical stability, defined as at least 4 weeks without antibiotics and/or oral corticosteroids to treat either a moderate or severe exacerbation - current or ex-smokers with a smoking history equivalent to at least 10 pack years (1 pack year = 20 cigarettes smoked per day for 1 year)	-having undergone major lung surgery (e.g., lung volume reduction, lung transplant) -having a lung tumour -primary respiratory diseases other than COPD (e.g., asthma) - impaired mobility related to non-COPD causes, as judged by the investigator
PD	-aged 18+ years -Diagnosis of PD according to the Movement Disorders Society criteria [23]	-impaired mobility related to non-PD causes, as judged by the investigator
MS	-aged 18+ years -Diagnosis of MS based on the revised McDonald's criteria	-impaired mobility related to non-MS causes, as judged by the investigator

PFF	-65+ years of age -surgical treatment (fixation or arthroplasty) for a low-energy fracture of the proximal femur (ICD*-10 diagnosis S72.0, S72.1, S72.2) as diagnosed on X-rays of the hip and pelvis within last 12 months	-impaired mobility related to non-PFF causes, as judged by the investigator
CHF	- ≥45 years of age -Diagnosis of chronic heart failure with a grading of II-IV of the New York Heart Association (NYHA) Classification	- history of COPD ≥GOLD III - impaired mobility related to non-CHF causes, as judged by the investigator
HA	65+ years of age	
	*ICD, International Classification of Disease	

Table 3 Generic and cohort-specific clinical outcomes.

Cohort	**Generic Outcomes**
All	- Descriptive measures (age, sex, living arrangements, education) - Anthropometric measures (height, mass, shoe size, waist width) - Health status (comorbidities, number of falls and injuries in the 12 months prior to assessment, walking aid usage and current medication) - Montreal Cognitive Assessment to evaluate global cognition [22] - Visual Analogue Scale to measure pain during walking (0-10, from no pain to worse pain possible) - Function component of the Late-Life Function and Disability Instrument (LLFDI) to evaluate function and disability [24,25]

Cohort	**Cohort-specific clinical outcomes**
PD	- Movement Disorder Society (MDS)-sponsored revision of the Unified Parkinson's Disease (PD) Rating Scale (UPDRS), motor part [26]
MS	- Expanded Disability Status Scale (EDSS)* [27]
COPD	- 6-minute walk test - Recent spirometry test obtained from medical notes to characterise lung function* - COPD Assessment Test™ (CAT) [28]
PFF	- Short Physical Performance Battery (SPPB) – quiet standing balance task, a five times chair-raise test, and a 4m walk test at preferred gait speed
CHF	- 6-minute walk test - Kansas City Cardiomyopathy Questionnaire (KCCQ) * [29]

* denotes measures obtained from medical records if completed within 6 months prior to assessment

3.2.3 Patient and Public Involvement

Patient and Public involvement and engagement (PPIE) has informed the design and conduct of this study. The protocol and patient facing documents were reviewed and changes implemented based upon reviewer feedback. We will work with our Patient Advisory Group (with members representing the patient cohorts involved in the study) to review study findings, data interpretation and study reporting and dissemination, including co-design and presentation of dissemination materials for patients and the public. The learnings derived from these activities will further inform the work of the wider Mobilise-D project.

3.2.4 Experimental protocol

Performance of algorithms to determine walking-related DMOs is mostly affected by three factors: (i) the type of motor task (e.g., slow as opposed to fast, straight as opposed to curvilinear or inclined walking, etc.), (ii) the population of interest (e.g., healthy vs pathological gait), (iii) the context of observation (e.g., home vs outdoors). To accommodate these factors, we have developed a comprehensive, multi-stage protocol that includes a variety of tests conducted both in a laboratory context and in the real world (Table 4).

Table 4 Summary of the experimental protocol used for the validation of the algorithms in the laboratory and in the real world. INDIP= INertial module with DIstance Sensors and Pressure insoles.

Context of assessment	Reference Systems	Tested device	Mobility Tasks
Laboratory	Stereophotogrammetry	DynaPort MM+	Structured mobility tasks and daily living activities
		INDIP	
Real world (2.5 hours)	INDIP	DynaPort MM+	Unsupervised real-world activities (including predetermined tasks)
	Mobile Phone with Aeqora App		
	Beacon		
Real world (7 days)	Mobile Phone with Aeqora App	DynaPort MM+	Unsupervised daily living
	Beacon		

a. Laboratory based assessment

Laboratory-based observations will be used to quantify validity and consistency within and between groups and different types of walking tasks under controlled ideal conditions. Structured and task-based mobility activities and a simulated daily activity session, mimicking habitual movements performed at home or at work will be included. The outcome of this comparison will provide the level of highest expected accuracy and minimum detectable changes for a given DMO.

Measurement tools

Reference system

A stereophotogrammetric (SP) system (100Hz) will be used as the gold standard in structured and simulated tests of daily activities to validate the DMOs calculated from the DynaPort MM+ raw data. SP systems provide a measurement of the instantaneous position of points in a 3D measurement volume, by means of a set of cameras, each of which can capture the 2D trajectories of markers that are attached to the object of interest. The trajectory reconstructions are affected by systematic and random instrumental errors, normally minimised to a few millimetres via ad hoc calibration procedures and filtering and smoothing techniques [30].

To ensure quality and consistency in the SP data collection, accommodating different SP systems across sites, a spot check designed following the methodology proposed by Di Marco et al. [31] will be used, which will establish the specific level of accuracy for each system. A graphical user interface (GUI) for automated pre-processing of the SP data will ensure consistency in associated procedures (labelling, gap filling, etc.).

A bespoke marker set will be adopted, including four markers on each foot for detecting the gait events and four markers on the lower back device to track the displacement of the DynaPort device (Figure 3).

[Figure 3 around here]

Tested devices

During this observation each participant will also be equipped with an additional multi-sensor system (INertial module with DIstance Sensors and Pressure insoles, INDIP) [32–34] and with the Dynaport MM+. The INDIP system (Figure 4) includes four inertial modules (one on the lower back, one on the non-dominant wrist and two on the feet), two distance sensors and two force-sensitive resistor pressure insoles including 16 force-resistive sensing elements (manufacturer 221e S.r.l., Italy). The INDIP has been designed to be used as a reference for real-world experiments, and in this phase of the protocol its performance will be validated against the SP system for the populations of interest. Spatio-temporal parameters will be estimated exploiting the sensors redundancy and implementing previously validated sensor fusion algorithms. Gait events will be detected using data from pressure insoles and inertial

sensors independently, and then combined to increase robustness and detection accuracy (missed and minimisation of extra events). Spatial variables will be computed from the inertial data of the feet using a Madgwick filter [35,36] combined with a zero-velocity update [37,38], subsequently velocity and displacement will be calculated using a direct and reverse approach [32,39–41]. The individual components of the INDIP system and the associated algorithms for the estimates of the DMOs have already been extensively validated in previous studies on various healthy and pathological cohorts [39,42,43]. The final assembled system in its fully synchronized configuration, developed to address the requirements of this study, is expected to perform equivalently and as such we can anticipate mean absolute percentage errors of 1% on the stride duration, between 2% and 3% in the estimate of the stride length. Preliminary results from in-lab validation showed percentage errors of about 2% for gait speed as estimated during continuous walking, including both straight and curvilinear portions [44].

[Figure 4 around here]

The lower back INDIP unit and the DynaPort MM+ will be rigidly attached to each other. The data from the SP (100Hz), INDIP (IMU and insoles, 100Hz, Distance sensor, 50Hz) and DynaPort MM+ (100Hz) systems will be synchronised using a hardware-based approach for the SP and the INDIP system, and timestamps to align recordings from the INDIP and the DynaPort MM+.

Mobility tasks

Structured Mobility Tasks

Straight Walking: Straight walking is the most common test of walking. [45,46] The participant walks for a distance of 5m from a standing start and will be repeated at three different walking speeds: preferred, fast and slow (Figure 5 A).

Timed Up and Go (TUG): The TUG is a widely used clinical assessment of a person's mobility [47]. The participant is asked to sit in a chair, stand up, walk 3 m in a straight line, make a 180° turn, walk back to the chair, turn and sit down (Figure 5 B).

L Test: The participant is asked to sit in a chair, stand up, walk straight, turn 90° to the left around a cone, walk straight to the second cone, make a 180° turn to the left, walk straight before making a final 90° turn to the right and return to the chair to sit down (Figure 5 C). Besides being a clinically validated test, [48] the main purpose of including this test is the variation in curvilinear walking and the inclusion of different types of turns.

Two novel additional tests were also included to simulate confounding factors that could be encountered in the real word:

Surface Test: The participant walks around a defined circuit by turning around the cones (Figure 5 D). The circuit is completed twice, creating the longest walking bout out of all the tasks (approx. 20m).

Hallway Test: The participant walks along a 6m walkway stepping up and down a step positioned in the walkway. At the end of the walkway, the participant will complete a sharp 180° turn and walk back

along the walkway (again stepping up and down off the step) until reaching the end point of the test (Figure 5 E).

[Figure 5 around here]

Daily living activities

These lab-based tasks will be used to simulate daily activities expected in the real life, similar to previous studies [49]. The participant starts by sitting in chair one and then executes a series of daily living tasks while moving around the room (see Figure 5F and G).

Patients will be given regular opportunities for rest periods and will be asked to communicate if they require any additional breaks or would like to stop the assessment at any point. Use of arm rests for the TUG, L-Test and SDA, as well as handrails for the hallway test are permitted when needed.

b. Real - world validation (2.5 hours observation)

This phase of the protocol will quantify validity and consistency across individuals and different types of walking tasks in the real world. It will be performed in a habitual environment (home/work/community/outdoor) chosen by the participants, without specific restrictions. The duration of the observation has been established as a trade-off between experimental, clinical and technical requirements.

Measurement tools

Reference systems

Participants will be asked to wear the INDIP, which in this phase of the protocol will be used as a reference system for the quantification of the DMOs provided by the single sensor algorithms, as applied to the DynaPort MM+ data.

In order to quantify the effects of contextual confounding factors, the participants will also be provided with a system detecting outdoors walking, gradient of descent/ascent (walking uphill/downhill). The system is developed as a mobile Android application (Aeqora app) and the device selected was a Samsung S9 with Android 10. The app is composed of three parts: *(i)* the core tracker, *(ii)* the interface and *(iii)* the server infrastructure collecting data across users. The core tracker, adapted from a library developed by the University of Sheffield [50], utilises the mobile phone's internal sensors to compute the type of activity (e.g., walking) and intensity (e.g., cadence) to identify geo-located bouts of movement. It operates in the background and senses mobility features through a range of sensors (e.g., step counters, activity recognition, accelerometer, gyroscope, etc.) as well as from location services (GPS, network, Bluetooth, etc.). It collects the data and stores the raw sensor data into a local database in real time. A set of mechanisms have been developed to control access to this data, keep it secure, and regulate its use. First, no user identity information is sent within a single request as a token identifier is

used. Additionally, a security layer is built based on Secure Socket Layer (SSL) and Transport Layer Security (TLS) 3.0 protocol to the data with scalable and efficient encryption algorithms. An SSL/TLS certificate is issued and used to establish identity and trust between server and client apps (desktop and mobile), ensuring privacy and security whenever communicating sensitive data.

Data collected during the experiments will be sent to a cluster of servers that uses algorithms to integrate the phone's data with contextual information about the locations where the participant will walk: where possible walks will be matched to OpenStreetMap [51] roads and paths, to remove GPS noise, the slope variation of each walk is computed on tiles, with a resolution of 5m within the UK (using Ordnance Survey Terrain 5) [52] and 30m in the other locations (using NASA's Shuttle Radar Topography Mission (SRTM) data [53], indoors and outdoors walking is recognised. Moreover, weather is associated with participant location based on the most proximate weather station.

The use of walking aids will also be monitored in this phase. For this purpose, a Bluetooth beacon (BlueBeacon Tag, BlueUp) will be attached to the walking aid and its activity will be detected by the phone's mobile tracker and saved by the app. The distance between the phone and the Beacon and data from the accelerometer contained in the Beacon will be integrated to determine when the aid is in use.

The above contextual factors and the use of walking aids will be included in the analyses to determine the extent to which they affect variation in the DMOs, although the degree of correlation will be adversely affected by the issues in accurately measuring context that are associated with missing data and GPS accuracy.

Mobility Tasks

To capture the largest possible range of activities during this assessment, participants will be guided by the following list of activities to be included: if relevant for their chosen environment, rise from a chair and walk to another room; walk to the kitchen and make a drink; walk up and down a set of stairs (if possible); walk outdoors (if possible, for a minimum of two minutes); if walking outside, walk up and down an inclined path. No supervision or structure to how these tasks should be completed will be given to the participants.

c. Real - world validation: seven days monitoring

This observation will quantify the effects of device wearing time, hourly and daily fluctuations of DMO's, and contextual confounding factors (such as location of the walk, weather, type of housing, etc.).

Measurement tools

The participants will be asked to wear the DynaPort MM+, and to carry a mobile phone equipped with the Aequora App. Bluetooth beacons will also be used to track the use of walking aids. The participants will wear the Dynaport MM+ at all times (including at night). As this device is not waterproof, they

will be instructed to remove it for showering, bathing, using a sauna and swimming and reattach it afterwards. They will be asked to keep the mobile phone charged, switched on at all times and to carry it with them whenever possible, especially when leaving the house.

Mobility Tasks

Participants will be monitored continuously for seven days, without any specific instruction being provided, except for that of wearing the provided measurement tools.

3.3 Assessment of participants' and assessors' experience

This part of the study will evaluate the participants' and assessors' experience of using the monitoring device. For the participant's assessment, wear-time of the device during the 7 days monitoring will be collected as a primary measure of compliance. Following the period of the seven-day, free-living data collection, participants will complete two questionnaires to assess the acceptability of the device. The first is a 12-item questionnaire [54] investigating usability on a 5-point ordinal scale. The questions are simple and focus on the impact of using a wearable device on participants' feelings, comfort and the ease of use of the device. The second questionnaire is the Comfort Rating Scale, [55] a 6-item measure investigating the comfort of a wearable device on a 21-point ordinal scale from '0 – low agreement' to '20 – high agreement'.

A subset of participants from all recruiting sites and cohorts will complete a semi-structured interview (see supplementary file). For this qualitative part of the study, sampling will continue until saturation, i.e. until no additional learning is identified from the data. The interview will explore participants' opinions on the use of wearable devices and digital technology in healthcare, experiences of managing their condition, experiences of technology, and opinions on data privacy associated with the use of technology in healthcare. Additionally, participants will be asked about their experiences of using the device, including comfort, perceived usefulness and ease of use, barriers and facilitators, and any other usability experiences that they may have encountered. All interviews will be audio-recorded, transcribed verbatim and, where required, translated to English.

To assess the professionals' experience, assessors from each of the clinical sites will be asked to assess the usability of the device after completion of the data collection. They will be provided with three questionnaires: 1) The System Usability Scale [56] a commonly used, validated 10-item questionnaire that asks users to rate a device on a 5-point Likert scale from '1 strongly disagree' to '5 – strongly agree'. Questions focus on the ease of use of the device, and the integration of various functions within it; 2) The IBM Computer System Usability Questionnaire [57] (to assess the DynaPort MM+ software), is composed of 19 items and asks respondents to consider their interaction with a computer system on a 7-point Likert scale from the perspective of data collection; 3) A bespoke questionnaire designed specifically for the TVS to assess the acceptability and effectiveness of the training methods, procedures, and any other materials provided within the study. The questionnaire will ask respondents to rate their experiences on a 7-point Likert scale to determine whether any changes to the procedures

and materials are required, and whether training was effective in preparing assessors to implement the assessment protocol as planned.

In addition, assessors will complete a semi-structured interview with the aim of exploring their experiences of the data collection process. Assessors will be asked about the use of the device (e.g., ease of use, intuitiveness, data collection and download procedures, etc.), training and materials provided prior to the commencement of the study, and barriers and facilitators to using the device. The topic guide and open-ended questions allow for new areas of conversation to emerge. All interviews will be audio-recorded and transcribed verbatim.

3.4 Data management

All data will be uploaded to a central platform "e-Science Central" (e-SC) [58] which provides data processing and storage functionality in accordance with principles of reproducible research. The underlying infrastructure complies with ISO 27000 standards for Information Security Management Systems and is hosted on Amazon Web Service secure services cloud platform. Data will be integrated on the platform by means of implementation of a standardised file nomenclature system. At point of capture, each file will be labelled in standardised format. For source data, we will adhere to principles defined by the US Food and Drug Administration [59] for making them attributable, legible, contemporaneous, original, and accurate (ALCOA+). In particular, we will use both web-based forms on e-SC and an application from ERT (partner in the project) to capture the electronic clinical outcome assessments (eCOAs). The e-SC forms provide storage of event data, and support for data validation and basic data entry and verification. Both e-SC and ERT systems employ error-handling at source which alert the assessors of incorrect data entry (e.g., min/max boundaries, required/optional fields). Data captured at source on paper will be copied, signed, and scanned, then uploaded to e-SC as a certified copy. The motion capture data will also be transferred to e-SC and stored in an unmodified form. These data will be either uploaded directly to e-SC via the e-SC portal or transferred via an Application Programming Interface (API). The algorithms being developed and benchmarked will be used to process these files and extract and store the DMOs.

3.5 Data analysis plan

3.5.1 Verification of the device

Mean and standard deviation readings of the accelerometer, gyroscope, and magnetometer signals captured during static acquisition will be used to assess the reliability of the manufacturer sensor calibration and to detect the presence of abnormal spikes in the sensor signals. Data from long static acquisitions will be used to confirm the stability of sampling frequency, the duration of the battery, and to estimate the Allan deviation (bias instability) of the gyroscope readings over time. Errors of

gyroscope readings will be assessed using mean and standard deviation of nominal, measured and relevant errors for angular velocity values during the dynamic acquisitions.

3.5.2 Validation of the algorithms

The data analysis will determine criterion validity (including selected performance metrics and criterion (concurrent) validity metrics of the primary (real world walking speed, RWS) and secondary DMOs listed in Table 5. Table 6 summarises the statistical tools that will be used to quantify each of the DMOs. All statistical analyses will be performed using the statistical analysis toolbox of Matlab R2018a. In all tasks and observations, continuous variables (e.g., cadence, real-walking-speed) will be summarised with descriptive statistics for the values obtained within walking bouts (mean and standard deviation). In addition, the mean, minimum, maximum, standard deviation, median, interquartile range (IQR) and Root Mean Square Error (RMSE) of DMOs over all available walking bouts will be presented. Categorical variables (e.g., laterality of initial contacts) will be summarised with frequency counts and percentages. Confidence Interval 95% (CI) will be provided for the Interclass Correlation Coefficients (ICC).

Using the gold standard as a reference, True Positives (TP), False Positives (FP), True Negatives (TN) and False Negatives (FN) will be identified for the DMOs identified from the single device using a cut-off tolerance window defined as a fixed interval of 0.5 seconds [60] and centred on each event detected by the reference system. The following performance metrics will then be calculated:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Positive\ Predicted\ Value = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Specificity = \frac{TN}{TN + FP} \text{ and}$$

$$F1\ score = 2 * \frac{Positive\ Predicted\ Value * Sensitivity}{Positive\ Predicted\ Value + Sensitivity}$$

Criterion validity will be characterised by evaluating the absolute and relative errors, defined as the relative and absolute differences between the DMOs quantified with the single device and those derived from the reference systems:

$$Relative\ error = \left(\left|\frac{DMO\ estimated\ by\ IMU - DMO\ estimated\ by\ Reference\ System}{DMO\ estimated\ by\ RS}\right|\right) \times 100$$

$$Absolute\ Error\ (DMO) = |DMO\ estimated\ by\ IMU - DMO\ estimated\ by\ Reference\ System|$$

The mean, standard deviation and maximum of all errors will be reported for each walking bout. Limits of agreement between single sensor and reference system DMOs will be quantified. In addition, statistically significant differences between the DMOs quantified by the IMU and those by the reference system, parametric (paired t-test) or non-parametric (Wilcoxon signed-rank test) tests will be performed depending on the normality of the distribution of the DMOs. Data distribution will be visually inspected with histograms, and normality tested with the Shapiro-Wilk test.

Concurrent validity between the DMOs quantified by the single device and those derived from the reference systems will be evaluated by quantifying the Intraclass Correlation Coefficient (ICC (2,1)).

All results will be presented separately by cohort (e.g., PD) and subgroup (i.e., subgroups of the cohorts) stratified by average stride gait speed (e.g., fast speed: walking speed > 1 m/s, medium speed: walking speed between 0.5m/s and 1 m/s, slow speed: walking speed < 0.5 m/s) [2].

If participants do not participate in one assessment (e.g., one of the tasks in the laboratory) or observation (e.g., 2.5hs), their remaining available data corresponding to remaining assessments/observations will still be included in the analyses. Within each of the contexts/assessments/observations, and assuming that data are missing completely at random, a complete case approach will be used to handle missing data [61].

Table 5 List of digital mobility outcomes (primary and secondary DMOs) that will be analysed as part of the TVS.

Variables	DMOs (units)	Definition	DMO Attainable			
			DynaPort MM+	SP System	INDIP	Aeqora App
Walking Bout (WB) *A walking sequence containing at least two consecutive strides of both feet. Start and end of a WB are determined by a resting period or any other activity (non-walking period).*	Number of WBs (count)	Based on the identification of gait as an activity (yes/no) to a sample level of 0.1 s	✓	✓	✓	✓
	WB Start (s)	Start of WB	✓	✓	✓	
	WB End (s)	End of WB	✓	✓	✓	
	WB Duration (s)	Time between start and the end of WB	✓	✓	✓	
Stride/ Step Duration (SD) *Refers to the duration (time intervals) of strides/steps, calculated as the time in between two non-consecutive (alternate) initial contacts.*	Stride Duration (s)	Duration between two non-consecutive (alternate) initial contact events	✓	✓	✓	
	Step Duration (s)	Duration between two consecutive initial contact events	✓	✓	✓	
Cadence (CE)	Cadence (steps/minute)	Steps performed within a minute	✓	✓	✓	
Stride Length (SL)	Mean Stride Length (m)	Average stride length within a WB	✓	✓	✓	
Real World Walking Speed (RWS)	Walking speed (m/s)	Velocity, average stride speed within a WB	✓	✓	✓	✓
Turning	Number of Turns	Overall number of turns performed in a WB based on the identification of turns (yes/no) to a sample level of 0.1 s	✓	✓	✓	
	Turn Start (s)	Start of each turn within the WB	✓	✓	✓	
	Turn End (s)	End of each turn within the WB	✓	✓	✓	
	Turn Duration (s)	Time between the start and the end of the turns within the WB	✓	✓	✓	
	Maximal Turn Angle (deg)	Maximal angle achieved in the turn	✓	✓	✓	

Height Estimation	Elevation Change (m)	Difference between the minimal and maximal height or elevation for the complete walking bout detected for incline walking	✓	✓	✓	✓
Left/Right Identification	Laterality (label)	Left or Right category, indicating the foot with which the initial contact is performed	✓	✓	✓	
Secondary Outcomes (SO)	Number of Final Contact Events (counts)	Correct identification of Final Contact events	✓	✓	✓	
	Final Contact Event (s)	Instant of time at which each final contact event is performed within a walking bout	✓	✓	✓	
	Swing Phase Duration (s)	Time between the last contact of the current footfall and the first contact of the next footfall on the same foot	✓	✓	✓	
	Stance Phase Duration (s)	Time in between the first contact and the last contact of two consecutive footfalls on the same foot	✓	✓	✓	
	Variability of: Step Time, Stride Time, Swing Time, Stance Time Stride Velocity Stride Length (same units as variable)	St. Dev. and Coefficient of Variation of step time, of stride time, of swing time, of stance time, of stride velocity and stride length within a WB	✓	✓	✓	
	Asymmetry of: Step Time, Stride Time, Swing Time, Stance Time (same units as variable)	Asymmetry evaluated as difference between right and left steps or strides for step time, of stride time, of swing time and of stance time within a WB	✓	✓	✓	
Contextual Factors	Location (label indoor/outdoor)	WB completed in an indoor or outdoor environment	✓			✓
	Walking Aid (label yes/no)	Walking aid assistance during WB	✓			✓

Table 6. List of statistical analyses and performance metrics that will be used for the various digital mobility outcomes (DMOs). Performance metrics and criterion validity are those that will be used to compare DMOs obtained from a single device versus those obtained from the reference system. The types of plots listed in the table will be used to visualise performance metrics and to support interpretation of the results.

DMO	PERFORMANCE METRIC					CRITERION VALIDITY							PLOTS		
	Sensitivity	Positive Predictive Value	Accuracy	Specificity	F1-score	Error (absolute & relative)	SD & Max of Error	Root Mean Squared Error	Precision	Concurrent validity (ICC)	Significant Difference	Limit of Agreement	Bland Altman Plots	Scatter Plots for Correlation	Histogram Plots
Number of Walking Bouts	✓	✓	✓	✓	✓										
WB Start						✓	✓	✓	✓				✓	✓	✓
WB End						✓	✓	✓	✓				✓	✓	✓
WB Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Initial Contact Events	✓	✓			✓	✓									
Step Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Stride Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Final Contact Events	✓	✓			✓	✓									
Mean Stride Length						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cadence						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Elevation Change						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Walking Speed						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Number of Turns	✓	✓	✓	✓	✓										
Turn Start						✓	✓	✓	✓				✓	✓	✓
Turn End						✓	✓	✓	✓				✓	✓	✓
Turn Duration						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Maximal Turn Angle						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Laterality						✓	✓	✓	✓	✓					
Sequence Order						✓	✓	✓							

3.5.3 Participants' and assessors' experience

Participant and professional questionnaire data will be analysed using descriptive statistics. Interviews will be analysed using thematic analysis. [62] Transcripts will be deductively examined for the presence of themes related to the acceptability of the monitoring device to participants (i.e., comfort, interference with daily living) based on previous literature. Specifically, perceived usefulness, comfort, and ease of use are critical factors of usability thus, these will be the categories examined within the transcripts. Regarding participants use of technology to manage their healthcare condition an inductive approach will be taken. A list of codes relevant to the question will be generated and then refined by grouping them into potential themes.

Authors' contributions:

This study is part of a large collaborative initiative, evolving around the design and execution of the study here described. The complexity of the protocol and the highly multi-disciplinary content of the

study justifies the high number of authors that have been and will be involved in the different stages of its planning, conducting and in the design of the analysis plan.

Manuscript initial drafting: CM, TB, SDD, AC, LA, AK, KS, SB, FC, LR. Design and deployment of experimental tools and protocols: KS, LS, CM, TB, EB, FC, NI, BS, LV, IV, EH, NC, FS. AC, MC, SB, LRO, AY, CK, EA, LS, CB, WM, CH, EW, JHA, EG, MB; J, MN, KT, BV, JHE, AK, DS, AM, VL. Data collection: LA, TB, PB, MB, EB, CF, EG, CH, EH, WM, CB, LP, LS, KS, BS, LVG, IV, AY, EW. Data analysis and algorithm development: SDD, EA, HS, AS, BE, FK, AK, MU, AC, JG, SK, LC, LP, LR, HH, EG. Draft revisions: LRO, KA, BE, JHA, BV, HS, AM, JHO, PB, FK, BC, AK, LC, LP, EB, KS, SK, LP, HH, WM, NC, BS, AI, IN, VL, SDD. Study design and coordination: CM, LRO, AM, SDD, AC. All authors have read and approved the final manuscript.

Competing interests statement

Fabio Ciravegna is CEO and shareholder of Aeqora Ltd, of which Vitaveska Lanfranchi is director and shareholder. Henrik Sillén is an employee of AstraZeneca. Bjoern M Eskofier is co-founder and owns shares of Portabiles HealthCare Technologies GmbH. McRoberts is the manufacturer of the DynaPort. Martijn Niessen, Jordi Evers, and Lucas Pluimgraaff are employees of McRoberts. Arne Mueller is an employee of Novartis and holds stock in Novartis. Lars Schwickert and Clemens Becker are consultants of Philipps Healthcare, Bosch Healthcare, Eli Lilly, Gait-up. Jeffrey M. Hausdorff reports having submitted a patent for assessment of mobility using wearable sensors in 400 Parkinson's disease; the intellectual property rights 401 are held by the Tel Aviv Medical Center. Luca Palmerini and Lorenzo Chiari are co-founders and own shares of mHealth Technologies.

Funding

This study was co-funded by the European Union's Horizon 2020 research and innovation programme and EFPIA via the Innovative Medicine Initiative 2 (Mobilise-D project, grant number IMI22017-13-7-820820). The views expressed are those of the authors and not necessarily those of the IMI, the European Union, the EFPIA, or any Associated Partners. We acknowledge the support of Grünenthal GmbH via the funding of a PhD scholarship directly dedicated to the technical validation protocol.

We acknowledge support from the Spanish Ministry of Science and Innovation through the "Centro de Excelencia Severo Ochoa 2019-2023" Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program.

SDD, AJY and LR are also supported by the Newcastle Biomedical Research Centre (BRC) based at Newcastle upon Tyne and Newcastle University. CM, BS, LVG and EB are also supported by the Sheffield Biomedical Research Centre (BRC) based at the Sheffield Teaching Hospital and the University of Sheffield. The work was also supported by the NIHR/Wellcome Trust Clinical Research Facility (CRF) infrastructure at Newcastle upon Tyne Hospitals NHS Foundation Trust and the CRF at the Sheffield Teaching Hospital. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care or the funders.

References

1 Team WHOD of MH and SA and WHO and WHOD of MH and SAMH and WHOMHE an. *Mental health atlas 2005*. World Health Organisation 2005.

2 Studenski S, Perera S, Patel K, *et al.* Gait Speed and Survival in Older Adults. *JAMA* 2011;**305**:50–8. doi:10.1001/jama.2010.1923

3 Perera S, Patel K V, Rosano C, *et al.* Gait Speed Predicts Incident Disability: A Pooled Analysis. *Journals Gerontol Ser A* 2016;**71**:63–71. doi:10.1093/gerona/glv126

4 Rochester L, Mazzà C, Mueller A, *et al.* A Roadmap to Inform Development, Validation and Approval of Digital Mobility Outcomes: The Mobilise-D Approach. *Digit Biomarkers* 2020;**4**:13–27. doi:10.1159/000512513

5 Del Din S, Godfrey A, Mazzà C, *et al.* Free-living monitoring of Parkinson's disease: Lessons from the field. *Mov Disord* 2016;**31**:1293–313. doi:10.1002/mds.26718

6 Bonci T, Keogh A, Din S Del, *et al.* An objective methodology for the selection of a device for continuous mobility assessment. *Sensors (Switzerland)* 2020;**20**:1–16. doi:10.3390/s20226509

7 Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. *Ind Manag Data Syst* 2015;**115**:1704–23. doi:10.1108/IMDS-03-2015-0087

8 Jia Y, Wang W, Wen D, *et al.* Perceived user preferences and usability evaluation of mainstream wearable devices for health monitoring. *PeerJ* 2018;**2018**:1–17. doi:10.7717/peerj.5350

9 Puri A, Kim B, Nguyen O, *et al.* User acceptance of wrist-worn activity trackers among community-dwelling older adults: Mixed method study. *JMIR mHealth uHealth* 2017;**5**. doi:10.2196/mhealth.8211

10 Viceconti M, Penna SH, Dartee W, *et al.* Toward a regulatory qualification of real-world mobility performance biomarkers in parkinson's patients using digital mobility outcomes. *Sensors (Switzerland)* 2020;**20**:1–13. doi:10.3390/s20205920

11 Mobilise-D website. https://www.mobilise-d.eu/ (accessed 18 Feb 2021).

12 Bartlett JW, Frost C. Reliability, repeatability and reproducibility: Analysis of measurement errors in continuous variables. *Ultrasound Obstet Gynecol* 2008;**31**:466–75. doi:10.1002/uog.5256

13 Letter of support for Mobilise-D digital mobility outcomes as monitoring biomarkers. https://www.ema.europa.eu/en/documents/other/letter-support-mobilise-d-digital-mobility-outcomes-monitoring-biomarkers_en.pdf (accessed 18 Feb 2021).

14 Systems M, Development S, Devices E. *IEEE Standard for Sensor Performance Parameter Definitions*. 2014.

15 El-Sheimy N, Hou H, Niu X. Analysis and modeling of inertial sensors using allan variance. *IEEE Trans Instrum Meas* 2008;**57**:140–9. doi:10.1109/TIM.2007.908635

16 Hussen AA, Jleta IN. Low-Cost Inertial Sensors Modeling Using Allan Variance. *Int Sch Sci Res Innov* 2015;**9**:1069–74.

17 IEEE Aerospace and Electronic Systems Society (institution). *IEEE Standard Specification Format Guide and Test Procedure for Single-Axis Laser Gyros*. 1995.

18 Ferraris F, Gorini I, Grimaldi U, *et al.* Calibration of three-axial rate gyros without angular velocity standards. *"Sensors Actuators, A Phys* 1994;**42**:446–9. doi:10.1016/0924-4247(94)80031-6

19 Stančin S, Tomažič S. Time- and computation-efficient calibration of MEMS 3D accelerometers and gyroscopes. *Sensors (Switzerland)* 2014;**14**:14885–915. doi:10.3390/s140814885

20 Kluge F, Del Din S, Cereatti A, *et al.* Consensus based framework for digital mobility monitoring. *medRxiv*. 2021;**1**;12 .doi:https://doi.org/10.1101/2020.12.18.20248404

21 COSMIN. https://www.cosmin.nl/ (accessed 18 Feb 2021).

22 Nasreddine ZS, Phillips NA, Bédirian V, *et al.* The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *J Am Geriatr Soc* 2005;**53**:695–9. doi:https://doi.org/10.1111/j.1532-5415.2005.53221.x

23 Postuma RB, Berg D, Stern M, *et al.* MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2015;**30**:1591–601. doi:10.1002/mds.26424

24 Jette AM, Haley SM, Coster WJ, *et al.* Late Life Function and Disability Instrument: I. Development and Evaluation of the Disability Component. *Journals Gerontol Ser A* 2002;**57**:M209–16. doi:10.1093/gerona/57.4.M209

25 Haley SM, Jette AM, Coster WJ, *et al.* Late Life Function and Disability Instrument: II. Development and Evaluation of the Function Component. *Journals Gerontol Ser A* 2002;**57**:M217–22. doi:10.1093/gerona/57.4.M217

26 Goetz CG, Tilley BC, Shaftman SR, *et al.* Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov Disord* 2008;**23**:2129–70. doi:https://doi.org/10.1002/mds.22340

27 Kurtzke JF. Rating neurologic impairment in multiple sclerosis. *Neurology* 1983;**33**:1444 LP – 1444. doi:10.1212/WNL.33.11.1444

28 Gupta N, Pinto LM, Morogan A, *et al.* The COPD assessment test: a systematic review. *Eur Respir J* 2014;**44**:873 LP – 884. doi:10.1183/09031936.00025214

29 Green CP, Porter CB, Bresnahan DR, *et al.* Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. *J Am Coll Cardiol* 2000;**35**:1245–55. doi:10.1016/S0735-1097(00)00531-3

30 Chiari L, Della Croce U, Leardini A, *et al.* Human movement analysis using stereophotogrammetry: Part 2: Instrumental errors. *Gait & posture* 2005; **21;**197-211. doi:10.1016/j.gaitpost.2004.04.004

31 Di Marco R, Rossi S, Castelli E, *et al.* Effects of the calibration procedure on the metrological performances of stereophotogrammetric systems for human movement analysis. *Meas J Int Meas Confed* 2017;**101**:265–71. doi:10.1016/j.measurement.2016.01.008

32 Trojaniello D, Cereatti A, Croce U Della. Gait direction of progression changes using shank worn MIMUs. In: *GBN Conference, Pavia*. 2014.

33 Salis F, Bertuletti S, Caruso M, *et al.* Multi-sensor integration and data fusion for enhancing gait assessment in and out of the laboratory. *Gait Posture* 2019;**74**:34. doi:10.1016/j.gaitpost.2019.07.493

34 Bertuletti S, Della Croce U, Cereatti A. A wearable solution for accurate step detection based on the direct measurement of the inter-foot distance. *J Biomech* 2019;**84**:274–7. doi:https://doi.org/10.1016/j.jbiomech.2018.12.039

35 Caruso, M.; Sabatini, A.M.; Laidig, D.; Seel, T.; Knaflitz, M.; Della Croce, U.; Cereatti, A. Analysis of the Accuracy of Ten Algorithms for Orientation Estimation Using Inertial and Magnetic Sensing under Optimal Conditions: One Size Does Not Fit All. *Sensors* 2021;**21**; 2543. doi:10.3390/s21072543

36 Caruso M, Sabatini AM, Knaflitz M, *et al.* Accuracy of the Orientation Estimate Obtained Using Four Sensor Fusion Filters Applied to Recordings of Magneto-Inertial Sensors Moving at Three Rotation Rates. In: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019;2053–8. doi:10.1109/EMBC.2019.8857655

37 Skog I, Nilsson J, Händel P. Evaluation of zero-velocity detectors for foot-mounted inertial navigation systems. In: *2010 International Conference on Indoor Positioning and Indoor Navigation*. 2010;1–6. doi:10.1109/IPIN.2010.5646936

38 Peruzzi A, Della Croce U, Cereatti A. Estimation of stride length in level walking using an inertial measurement unit attached to the foot: A validation of the zero velocity assumption during stance. *J Biomech* 2011;**44**:1991–4. doi:10.1016/j.jbiomech.2011.04.035

39 Trojaniello D, Cereatti A, Pelosin E, *et al.* Estimation of step-by-step spatio-temporal parameters of normal and impaired gait using shank-mounted magneto-inertial sensors : application to elderly , hemiparetic , parkinsonian and choreic gait. 2014;**44**:152.

40 Köse A, Cereatti A, Della Croce U. Bilateral step length estimation using a single inertial measurement unit attached to the pelvis. *J Neuroeng Rehabil* 2012;**9**. doi:10.1186/1743-0003-9-9

41 Trojaniello D, Cereatti A, Ravaschio A, *et al.* Assessment of gait direction changes during straight-ahead walking in healthy elderly and Huntington Disease patients using a shank worn MIMU. In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. 2014. 2508–11. doi:10.1109/EMBC.2014.6944132

42 Betoli M, Cereatti A, Trojaniello D, *et al*. Estimation of spatio-temporal parameters of gait from magneto-inertial measurement units: multicenter validation among Parkinson, mildly cognitively impaired and healthy older adults. *Biomedical engineering online.* 2018 Dec;**17**(1):1-4. doi:10.1186/s12938-018-0488-2

43 Rossaningo R, Caruso M, Salis F, *et al*. An Optimal Procedure for Stride Length Estimation Using Foot-Mounted Magneto-Inertial Measurement Units. *2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA)* 2021 Jun 23 (pp. 1-6). IEEE., doi: 10.1109/MeMeA52024.2021.9478604.

44 Salis F, Bertuletti S, Caruso M, *et al.* A novel multi-sensor system for gait assessment in real-world conditions: preliminary results. In: *Proceedings of GNB 2020*. 2020.

45 Zijlstra W, Hof AL. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture* 2003;**18**:1–10. doi:https://doi.org/10.1016/S0966-6362(02)00190-X

46 Ngueleu AM, Blanchette AK, Bouyer L, *et al.* Design and accuracy of an instrumented insole using pressure sensors for step count. *Sensors (Switzerland)* 2019;**19**:12–4. doi:10.3390/s19050984

47 Nightingale CJ, Mitchell SN, Butterfield SA. Validation of the Timed Up and Go Test for Assessing Balance Variables in Adults Aged 65 and Older. *J Aging Phys Act*;**27**:230–3. doi:10.1123/japa.2018-0049 10.1123/japa.2018-0049

48 Deathe AB, Miller WC. The L Test of Functional Mobility: Measurement Properties of a

Modified Version of the Timed “Up & Go” Test Designed for People With Lower-Limb Amputations. *Phys Ther* 2005;**85**:626–35. doi:10.1093/ptj/85.7.626

49 Pham MH, Elshehabi M, Haertner L, *et al.* Validation of a Step Detection Algorithm during Straight Walking and Turning in Patients with Parkinson’s Disease and Older Adults Using an Inertial Measurement Unit at the Lower Back . Front. Neurol. . 2017;**8**:457.https://www.frontiersin.org/article/10.3389/fneur.2017.00457

50 Ciravegna F, Gao J, Ireson N, *et al.* Active 10: Brisk Walking to Support Regular Physical Activity. In: *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. New York, NY, USA: : Association for Computing Machinery 2019. 11–20. doi:10.1145/3329189.3329208

51 Open Street Map. https://www.openstreetmap.org (accessed 18 Feb 2021).

52 Ordnance Survey Terrain 5. https://www.ordnancesurvey.co.uk/business-government/products/terrain-5 (accessed 18 Feb 2021).

53 U.S. Releases Enhanced Shuttle Land Elevation Data. https://www2.jpl.nasa.gov/srtm/ (accessed 18 Feb 2021).

54 Rabinovich RA, Louvaris Z, Raste Y, *et al.* Validity of physical activity monitors during daily life in patients with COPD. *Eur Respir J* 2013;**42**:1205–15. doi:10.1183/09031936.00134312

55 Knight JF, Baber C, Schwirtz A, *et al.* The comfort assessment of wearable computers. *Proc - Int Symp Wearable Comput ISWC* 2002;**2002-Janua**:65–72. doi:10.1109/ISWC.2002.1167220

56 Brooke J. SUS: a “quick and dirty’usability". *Usability Eval Ind* 1196;**Jun 11**:189.

57 Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. *Int J Human–Computer Interact* 1995;**7**:57–78. doi:10.1080/10447319509526110

58 Hiden H, Woodman S, Watson P, *et al.* Developing cloud applications using the e-Science Central platform. *Philos Trans R Soc A Math Phys Eng Sci* 2013;**371**. doi:10.1098/rsta.2012.0085

59 Data Integrity and Compliance With CGMP. Guidance for Industry. https://www.fda.gov/files/drugs/published/Data-Integrity-and-Compliance-With-Current-Good-Manufacturing-Practice-Guidance-for-Industry.pdf (accessed 18 Feb 2021).

60 Tietsch M, Muaremi A, Clay I, *et al*. Robust Step Detection from Different Waist-Worn Sensor Positions: Implications for Clinical Studies. *Digital biomarkers* 2020;4(1):50-8. doi:10.1159/000511611

61 Kang H. The prevention and handling of the missing data. *Korean journal of anesthesiology*.

2013 May;64(5):402. doi:10.4097/kjae.2013.64.5.402

62 Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;**3**:77–101. doi:10.1191/1478088706qp063oa

Figure legends

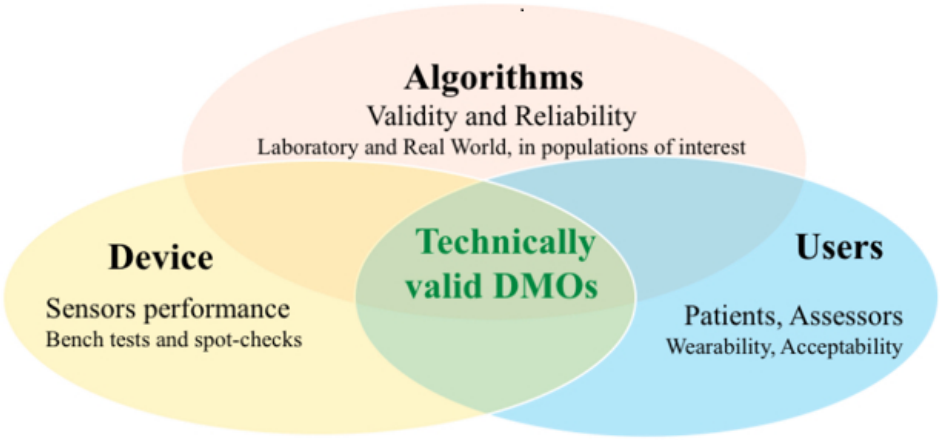
Figure 1 – Concurrent domains to be assessed as part of a technical validation of digital mobility outcomes (DMOs) obtained from a wearable device.

Figure 2 – Testing configurations used during (a) short static (plexiglass cube and device) and (b) dynamic (turntable and device) acquisitions. Marks (in red) are applied on both on the turntable and on the base to identify start/end, which have to align for the dynamic tests.

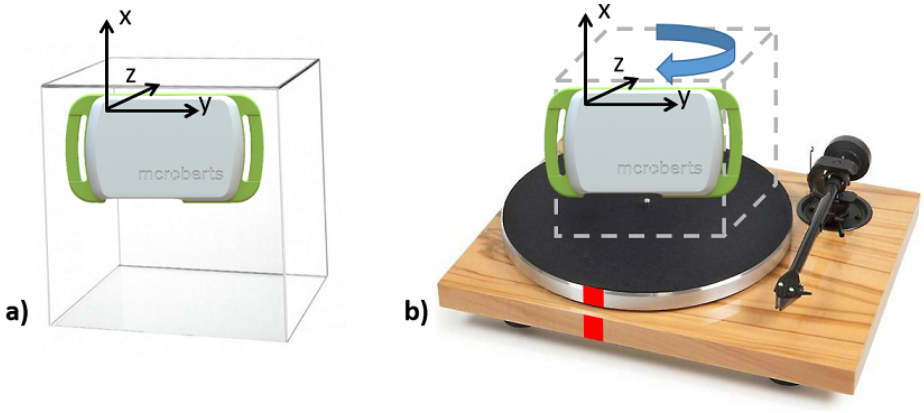
Figure 3 – Illustration of the adopted marker set configuration. Markers were located on the right (RHEEL) and left (LHEEL) heels, toes (RTOE, LTOE) and on the INDIP units located on the right and left foot (RINDIP, LINDIP). Two additional reference markers were asymmetrically attached to the side of the foot to favour automatic recognition (RREF. LREF). Four additional markers were located on the DynaPort MM+ sensor (DYNAY, DYNAO, DYNAX, DYNAREF).

Figure 4 –Different components of the INDIP system. The figure on the left shows the pressure insoles and the connectors that link them to the distance sensors and the inertial modules. The picture on the right shows how the same system is then attached to the participant's foot and leg.

Figure 5 – Diagrams of the selected tasks: (A) Straight Walking Test, (B) Timed Up and Go, (C) L-Test, (D) Surface Test, (E) Hallway Test, F) Schematic of the Daily living activities, G) Description of the eight tasks performed during the Daily Living activities.



Concurrent domains to be assessed as part of a technical validation of digital mobility outcomes (DMOs) obtained from a wearable device.



Testing configurations used during (a) short static (plexiglass cube and device) and (b) dynamic (turntable and device) acquisitions. Marks (in red) are applied on both on the turntable and on the base to identify start/end, which have to align for the dynamic tests.

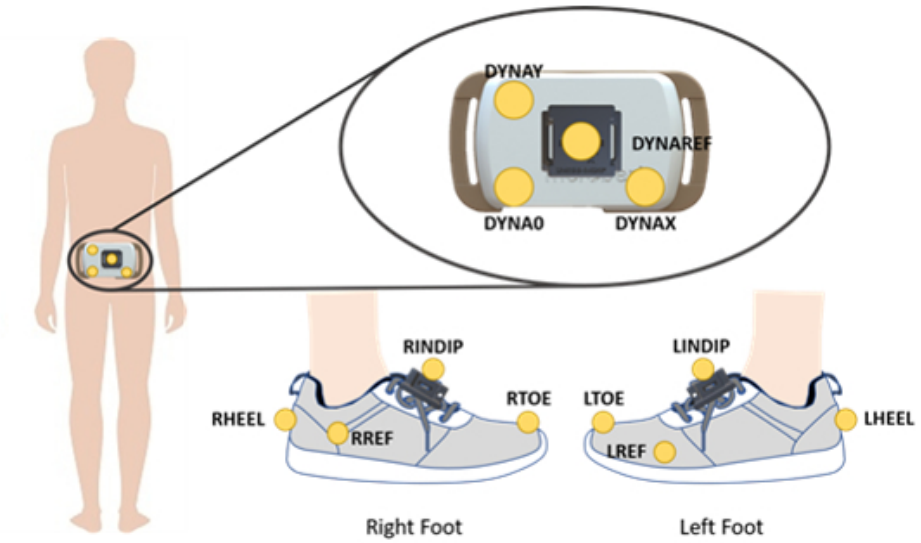
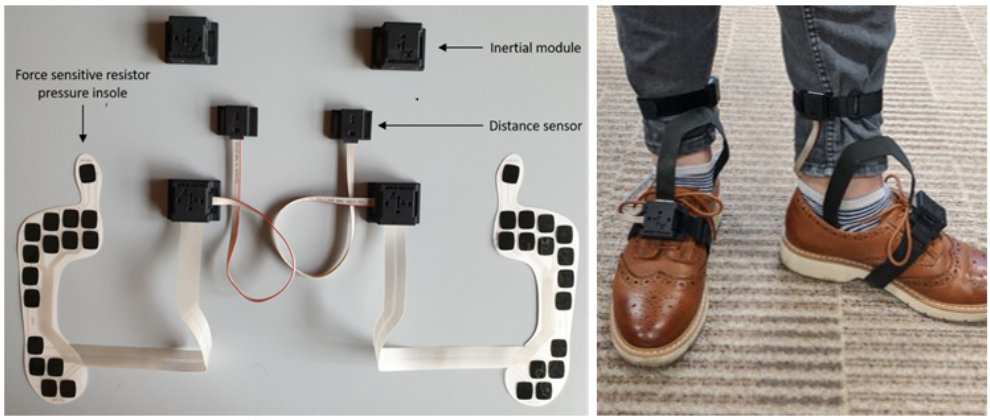
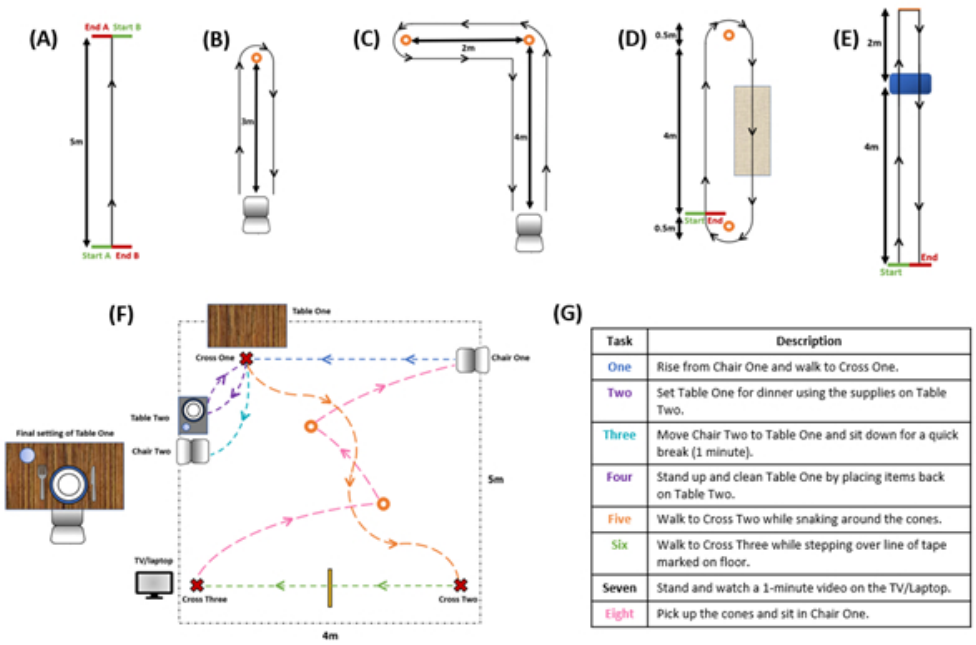


Illustration of the adopted marker set configuration. Markers were located on the right (RHEEL) and left (LHEEL) heels, toes (RTOE, LTOE) and on the INDIP units located on the right and left foot (RINDIP, LINDIP). Two additional reference markers were asymmetrically attached to the side of the foot to favour automatic recognition (RREF. LREF). Four additional markers were located on the DynaPort MM+ sensor (DYNAY, DYNAO, DYNAX, DYNAREF).



Different components of the INDIP system. The figure on the left shows the pressure insoles and the connectors that link them to the distance sensors and the inertial modules. The picture on the right shows how the same system is then attached to the participant's foot and leg.



(G)

Task	Description
One	Rise from Chair One and walk to Cross One.
Two	Set Table One for dinner using the supplies on Table Two.
Three	Move Chair Two to Table One and sit down for a quick break (1 minute).
Four	Stand up and clean Table One by placing items back on Table Two.
Five	Walk to Cross Two while snaking around the cones.
Six	Walk to Cross Three while stepping over line of tape marked on floor.
Seven	Stand and watch a 1-minute video on the TV/Laptop.
Eight	Pick up the cones and sit in Chair One.

Diagrams of the selected tasks: (A) Straight Walking Test, (B) Timed Up and Go, (C) L-Test, (D) Surface Test, (E) Hallway Test, F) Schematic of the Daily living activities, G) Description of the eight tasks performed during the Daily Living activities.

159x102mm (96 x 96 DPI)

Mobilise-D Technical Validation Study

Participant interview topic guide

Interview guide

Note: Questions in this topic guide are included to answer the above aims. If the participant is open and talks freely, they may answer some of the questions without being asked. Therefore, depending on the person, not all of these questions need to be asked. If they begin to talk about topics that may be interesting or relevant to the above aims, please feel free to continue to explore these, even if there is no specific question linked to it. In contrast, if participants are not very open, some potential prompts have been included with the questions below. These prompts are there as an optional guide and do not need to be used.

Dynaport questions	
Aim: Ensure that the McRoberts Dynaport device is comfortable and acceptable to participants	
Main questions	
1	Can you describe your experience of using the Dynaport sensor in the last week?
2	Can you tell me what you liked about the device? Disliked? *Prompts if needed* - Size/weight - Attachment to body - Ease of use - Comfort
3	How did the device make you feel? / Can you describe what it felt like to wear the device? *Prompts if needed* - In social environments/at home - Interaction with daily activities - Emotions associated with being monitored
4	What you change about the device if you could?
Follow-up questions **"**If you don't mind, I'd like to ask you some specific details about the device."	
1	How did you find the process of putting it on and taking it off?
2	How did the device influence your daily activities? *Prompts if needed* - How were they impacted?

- How did this make them feel?

3 Can you tell me about any difficulties that you had with the device?

4 How did you feel about wearing the device for a week?

Prompts if needed

- How would they feel if it was longer?
- Any concerns for the week?

Closing question

1 Is there anything else you would like me to know about the Dynaport?

The use of wearable devices in healthcare questions

Aim: Explore the acceptability of wearable devices, for the purposes of healthcare monitoring, to participants in general.

Main questions

1 Can you tell me about your experience of your health condition?

Note: Condition specific symptoms are listed at the end of this document.

2 Can you tell me about your experience of the care you've received for your condition?

Prompts if needed

- How do they feel about it?

3 Can you tell me about what sort of technology you currently use in your everyday life?

Prompts if needed

- How do you feel about using technology?
- What would make you use technology more? Less?
- Emotions associated with being monitored

4 What are your opinions on the use of technology in healthcare?

Prompts if needed

- What do you think it can be used for?

5 How would you feel about using technology to generate health information about yourself? (e.g. condition related smartphone app, self-reported outcomes platform, fitness tracker etc.),

Prompts if needed

- Why?
- What would make you use it?
- What would stop you from using it?

- What would need to change for you to use it?

Follow-up questions	
1	How do you feel about capturing health information in your daily life, using a wearable remote monitoring device?
2	How do you think digital technology used in your daily life would influence how you manage your condition? *Prompts if needed* - How would it impact their relationship with their health care provider? - Integration into activities of daily living
3	How would you feel about sharing this data with your health care provider? What about researchers? *Prompts if needed* - Usefulness - Impact of this

Closing question	
1	Is there anything else you would like me to know about using technology in healthcare?

Researcher feedback

Questionnaires

The System Usability Scale

To be completed by researchers who have collected data using the McRoberts device. When answering these questions please consider how you personally have found the McRoberts device, not how the participants have found it.

	1 strongly disagree	2	3	4	5 strongly agree
I think that I would like to use this system frequently.					
I found the system unnecessarily complex.					
I thought the system was easy to use.					

I think that I would need the support of a technical person to be able to use this system.					
I found the various functions in this system were well integrated.					
I thought there was too much inconsistency in this system.					
I would imagine that most people would learn to use this system very quickly.					
I found the system very cumbersome to use.					
I felt very confident using the system.					
I needed to learn a lot of things before I could get going with this system.					
I think that I would like to use this system frequently.					

The IBM Computer System Usability Questionnaire

To be completed by researchers who have collected data using the McRoberts device. When answering these questions please consider how you personally have found the McRoberts device, not how the participants have found it.

	1 strongly disagree	2	3	4	5	6	7 Strongly agree
Overall, I am satisfied with how easy it is to use this system.							
It is simple to use this system.							
I can effectively complete my work using this system.							

I am able to complete my work quickly using this system.							
I am able to efficiently complete my work using this system.							
I feel comfortable using this system.							
It was easy to learn to use this system.							
I believe I became productive quickly using this system.							
The system gives error messages that clearly tell me how to fix problems.							
Whenever I make a mistake using the system, I recover easily and quickly.							
The information (such as on-line help, on-screen messages and other documentation) provided with this system is clear.							
It is easy to find the information I need.							
The information provided with the system is easy to understand.							
The information is effective in helping me complete my work.							
The organization of information on the system screens is clear.							
The interface of this system is pleasant.							
I like using the interface of this system.							
This system has all the functions and capabilities I expect it to have.							

Intervention specific questionnaire

This questionnaire aims to evaluate your experiences and opinions on the processes of the Mobilise-D technical validation study. In particular, we are looking to focus on the components of the study which will also take place within the clinical validation trial. Therefore, when asking about certain aspects of the trial, we may be specific in relation to which device or measurement tool we want

feedback on. Please read each question carefully to ensure that your answers relate to the specific component under investigation.

The purpose of this questionnaire is to determine your opinions on i) the training you received, ii) the supporting materials that were provided to you, iii) the feasibility of conducting participant recruitment. Please answer each question honestly. All responses will be de-identified and treated anonymously.

Training

When considering training, we want you to consider the training that you received to carry out the following:

- Recruit participants
- Complete the human factors assessment
- Use the McRoberts device
- Use the ERT platform
- Use the EScience platform

Many of these training components would have been delivered separately. We want you to consider your experience on these training sessions overall, and then answer questions for each specific component independently. There is space to write comments at the end of this section if you wish to add more information or clarify any aspect further.

Overall training	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
The Mobilise-D training programme was enjoyable						
The Mobilise-D training programme was useful						
The Mobilise-D training programme successfully prepared me to collect data within the technical validation trial						

The Mobilise-D training programme successfully prepared me to recruit participants within the technical validation trial						
Overall training content	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
The training was sufficiently interactive						
The training was interesting						
The training was easy to understand						
The training provided a clear outline of what was expected from me						
The training provided a clear rationale of recruitment processes						
Recruitment training	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
I felt confident in my recruitment role following training						
I felt competent to complete recruitment following training						
Following training, I had no questions regarding recruitment						
Following training I was confident in who I should						

contact if I had a question regarding recruitment						
Data collection training	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
I felt confident in my data collection role following training						
I felt competent to complete data collection following training						
Following training, I had no questions regarding data collection						
Following training I was confident in who I should contact if I had a question regarding data collection						
McRoberts training	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
I felt confident in using McRoberts following training						
I felt competent to complete data collection with McRoberts following training						
Following training, I had no questions regarding the McRoberts device						
Following training I was confident in who I should contact if I had a question regarding McRoberts						

ERT platform training	Strongly disagree 1	Disagree 2	Neutral 3	Agree 4	Strongly agree 5	Not applicable
I felt confident in the use of the ERT platform following training						
I felt competent to complete data collection using ERT following training						
Following training, I had no questions regarding the ERT platform						
Following training I was confident in who I should contact if I had a question regarding the ERT platform						

Please provide us with any further feedback you have regarding the training that you undertook as part of the Mobilise-D project:

Project materials feedback

As part of the Mobilise-D project, you received a number of materials to help support you in the recruitment and data collection processes (i.e. manuals, etc). Please provide us with your feedback on these items

Recruitment materials	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
The recruitment materials were easy to understand						
I used the recruitment materials a lot to help me recruit participants						
As recruitment progressed I used the materials less						
The materials answered all my recruitment questions						
The materials made the recruitment process easier						
When I used the materials I felt more confident in the recruitment process						
Data collection materials	**Strongly disagree 1**	**Disagree 2**	**Neutral 3**	**Agree 4**	**Strongly agree 5**	**Not applicable**
The data collection materials were easy to understand						
I used the data collection materials a lot to help me recruit participants						
As data collection progressed I used the materials less						
The materials answered all my data collection questions						

The materials made the data collection process easier
When I used the materials I felt more confident in the data collection process

Please provide us with any further feedback you have regarding the materials that you received as part of the Mobilise-D project:

_____

_____

_____

_____

_____

_____

_____

Feasibility of the trial procedures

Recruitment

How many participants did your site recruit?

How many participants was your site due to recruit?

What barriers to participant recruitment did you encounter?

What helped you to recruit people?

What additional materials do you believe would help support recruitment in the future?

What additional support do you believe would help support recruitment in the future?

Data collection

How many participants failed to complete full data collection procedures?

Why?

What additional materials do you believe would help support data collection in the future?

What additional support do you believe would help support data collection in the future?

Do you have any other thoughts regarding the recruitment and data collection procedures of the Mobilise-D technical validation trial?

Researcher feedback

Interview guide

Start with a general question about how they found the experience of data collection within the Mobilise-D validation study

The Dynaport device

- Tell me about how you found using the Dynaport sensor during the trial?
- What were your first impressions of the device?
- Can you tell me what you liked about the device? Disliked?
- Can you tell me about any difficulties that you had with the device or its platform during the trial?
    - Difficulties they encountered themselves
        - Set up
        - Ease of explanation for participants
    - Difficulties that were reported to them
        - Did participant get in contact during the week?
        - Did any devices come back damaged?
- How would you compare this device to other wearables that you have used before?

The ERT device

- Tell me about how you found using the ERT platform during the trial?
- What were your first impressions of the device?
- Did any of your opinions change as the trial progressed?
- Can you tell me what you liked about the device? Disliked?
- Can you tell me about any difficulties that you had with the device during the trial?
- How did the device make you feel?

The EScience platform

- Tell me about how you found using the EScience platform during the trial?

- What were your first impressions of the platform?
- Did any of your opinions change as the trial progressed?
- Can you tell me what you liked about the platform? Disliked?
- Can you tell me about any difficulties that you had with the device during the trial?

The training and materials used

- Can you tell me what your opinions are of the training you received before starting data collection?

  Specifically related to: Dynaport, EScience, ERT and recruitment procedures
  - Duration
  - Content
  - Facilitators
  - Materials provided
- What did you expect to get out of the training?
- What was useful about the training? Not useful?
- Describe how you found the process of data collection?
- How did this compare to what you expected following training?
- What would you change about the training if you could? And the materials provided?

Recruitment procedures

Can you tell me about your experiences recruiting participants for the Mobilise-D trial
What needs to change to make recruitment easier?