

BMJ Open Technical validation of real-world monitoring of gait: a multicentric observational study

Claudia Mazzà ^{1,2}, Lisa Alcock,³ Kamiar Aminian,⁴ Clemens Becker,⁵ Stefano Bertuletti,⁶ Tecla Bonci,^{1,2} Philip Brown,⁷ Marina Brozgol,⁸ Ellen Buckley ^{1,2}, Anne-Elie Carsin ^{9,10,11,12}, Marco Caruso,^{13,14} Brian Caulfield,^{15,16} Andrea Cereatti,¹³ Lorenzo Chiari,^{17,18} Nikolaos Chynkiamis,¹⁹ Fabio Ciravegna,^{1,20} Silvia Del Din,³ Björn Eskofier,²¹ Jordi Evers,²² Judith Garcia Aymerich,^{9,10,11} Eran Gazit,⁸ Clint Hansen ²³, Jeffrey M Hausdorff,^{8,24} Jorunn L Helbostad,²⁵ Hugo Hiden,²⁶ Emily Hume,¹⁹ Anisoara Paraschiv-Ionescu,⁴ Neil Ireson,^{1,20} Alison Keogh,^{15,16} Cameron Kirk,³ Felix Kluge,²¹ Sarah Koch,^{9,10,11} Arne Küderle,²¹ Vitaveska Lanfranchi,^{1,20} Walter Maetzler,²³ M Encarna Micó-Amigo,³ Arne Mueller,²⁷ Isabel Neatrou,³ Martijn Niessen,²² Luca Palmerini,^{17,18} Lucas Pluimgraaff,²² Luca Reggi,¹⁸ Francesca Salis,⁶ Lars Schwickert,⁵ Kirsty Scott,^{1,2} Basil Sharrack,²⁸ Henrik Sillen,²⁹ David Singleton,^{15,16} Abolfazi Soltani,⁴ Kristin Taraldsen,²⁵ Martin Ullrich,²¹ Linda Van Gelder,^{1,2} Beatrix Vereijken,²⁵ Ioannis Vogiatzis,¹⁹ Elke Warmerdam,²³ Alison Yarnall,^{3,7} Lynn Rochester^{3,7}

To cite: Mazzà C, Alcock L, Aminian K, *et al*. Technical validation of real-world monitoring of gait: a multicentric observational study. *BMJ Open* 2021;**11**:e050785. doi:10.1136/bmjopen-2021-050785

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2021-050785>).

Received 02 March 2021
Accepted 28 October 2021



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to

Professor Claudia Mazzà;
c.mazza@sheffield.ac.uk

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, a consortium funded by the European Union (EU) as 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 with 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.

Strengths and limitations of this study

- A multidisciplinary 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 multistage and multidevice 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.

Ethics and dissemination The study 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.

Trial registration number ISRCTN (12246987).

INTRODUCTION

The ability to move is a key contributor to physical, mental and social well-being, which is in line with the 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 a 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 evaluation⁶ of the participants perception and acceptability of the device⁷ 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 clinical⁴ and regulatory¹⁰ purposes (see [figure 1](#) for summary).

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 multistage protocol aims to: (1) verify the metrological performance of the sensors included in a IMU-based monitoring device, using a procedure that could be replicated on any device; (2) establish the validity and reliability¹² of the DMOs estimated by the algorithms using data from an IMU-based device, taking into accounts the effects of populations (eg, healthy adults, patients with various conditions), locomotor activities (simple straight walking vs complex walking tasks), contexts (lab based vs 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 (3) 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, multimodal solution to validate real-world mobility. Notably, the protocol described here has been accepted by the European Medical Agency¹³ as part of the Mobilise-D process for the regulatory qualification of real-world mobility performance biomarkers in Parkinson's disease.¹⁰

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.

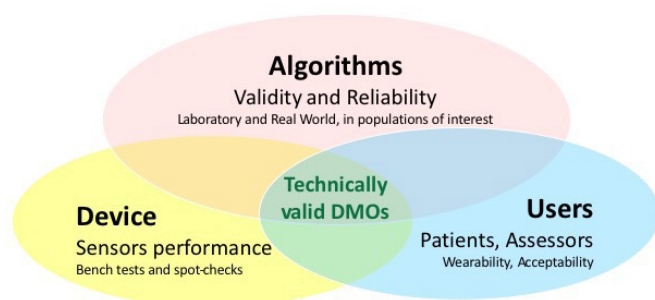


Figure 1 Concurrent domains to be assessed as part of a technical validation of digital mobility outcomes (DMOs) obtained from a wearable device.

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

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

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: 106.6×58×11.5 mm, size: 55 g).

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

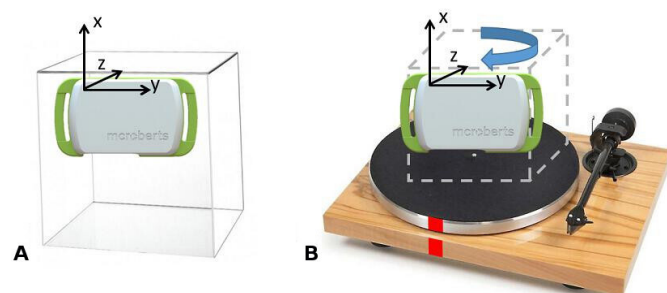


Figure 2 Testing configurations used during (A) short static (plexiglass cube and device) and (B) dynamic (turntable and device) acquisitions. Marks (in red) are applied on both on the turntable and on the base to identify start/end, which have to align for the dynamic tests.

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°/s) (figure 2).

The tests described above will be performed on 35 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.

Protocol for validation of the algorithms

Several algorithms to detect the DMOs from a single device have been implemented according to agreed definitions²⁰ 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.

Ethics and dissemination

This multicentre study is sponsored and coordinated by The Newcastle upon Tyne Hospitals National Health Service (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 2020 and last 6 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.

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 HA. Twenty participants will be recruited for each cohort. Each cohort will be

recruited across multiple sites to ensure generalisability (eg, 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.²¹ This sample size, however, will be refined after 50% of the data collection. Given that 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 Intraclass Correlation Coefficient (ICC) ≥ 0.7 , with Alpha=0.05 and Beta=0.9, and an aimed CI of 0.1. Based on this review, more participants may be recruited.

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

Group	Inclusion criteria	Exclusion criteria
All groups	<ul style="list-style-type: none"> ▶ Able to walk 4 m independently with or without walking aids ▶ Able to give informed consent ▶ Willingness to wear the sensor setups during the study ▶ Shoe size 36 European Union (EU) (3 UK) or above ▶ Able to read and write in first language of the respective country ▶ Montreal Cognitive Assessment >15⁵⁵ ▶ Available for home /office visit during study period 	<ul style="list-style-type: none"> ▶ 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	<ul style="list-style-type: none"> ▶ ≥ 45 years of age ▶ Diagnosis of COPD (post-bronchodilator forced expiratory volume in the first second (FEV₁) to forced vital capacity 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) 	<ul style="list-style-type: none"> ▶ Having undergone major lung surgery (eg, lung volume reduction, lung transplant) ▶ Having a lung tumour ▶ Primary respiratory diseases other than COPD (eg, asthma) ▶ Impaired mobility related to non-COPD causes, as judged by the investigator
PD	<ul style="list-style-type: none"> ▶ Aged 18+ years ▶ Diagnosis of PD according to the Movement Disorders Society criteria⁵⁶ 	<ul style="list-style-type: none"> ▶ Impaired mobility related to non-PD causes, as judged by the investigator
MS	<ul style="list-style-type: none"> ▶ Aged 18+ years ▶ Diagnosis of MS based on the revised McDonald's criteria 	<ul style="list-style-type: none"> ▶ Impaired mobility related to non-MS causes, as judged by the investigator
PFF	<ul style="list-style-type: none"> ▶ 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 	<ul style="list-style-type: none"> ▶ Impaired mobility related to non-PFF causes, as judged by the investigator
CHF	<ul style="list-style-type: none"> ▶ ≥ 45 years of age ▶ Diagnosis of chronic heart failure with a grading of II-IV of the New York Heart Association Classification 	<ul style="list-style-type: none"> ▶ History of COPD \geq GOLD III ▶ Impaired mobility related to non-CHF causes, as judged by the investigator
HA	65+ years of age	

CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; HA, healthy older adults; ICD, International Classification of Disease; MS, multiple sclerosis; PD, Parkinson's disease; PFF, proximal femoral fracture.

Table 3 Generic and cohort-specific clinical outcomes

Cohort	Generic outcomes
All	<ul style="list-style-type: none"> ▶ 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 cognition⁵⁵ ▶ 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 to evaluate function and disability^{57 58}
Cohort	Cohort-specific clinical outcomes
PD	▶ Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease (PD) Rating Scale, motor part ⁵⁹
MS	▶ Expanded Disability Status Scale ⁶⁰
COPD	<ul style="list-style-type: none"> ▶ 6 min walk test ▶ Recent spirometry test obtained from medical notes to characterise lung function* ▶ COPD Assessment Test⁶¹
PFF	▶ Short Physical Performance Battery—quiet standing balance task, a five times chair-raise test, and a 4 m walk test at preferred gait speed
CHF	<ul style="list-style-type: none"> ▶ 6 min walk test ▶ Kansas City Cardiomyopathy Questionnaire⁶²

*Denotes measures obtained from medical records if completed within 6 months prior to assessment.
 CHF, congestive heart failure; COPD, chronic obstructive pulmonary disease; MS, multiple sclerosis; PFF, proximal femoral fracture.

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](#).

Patient and public involvement

Patient and public involvement and engagement has informed the design and conduct of this study. The protocol and patient facing documents were reviewed and changes implemented based on 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 codesign 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.

Experimental protocol

Performance of algorithms to determine walking-related DMOs is mostly affected by three factors: (1) the type of motor task (eg, slow as opposed to fast, straight as opposed to curvilinear or inclined walking, etc), (2) the population of interest (eg, healthy vs pathological gait), (3) the context of observation (eg, home vs outdoors). To accommodate these factors, we have developed a comprehensive, multistage protocol that includes a variety of tests conducted both in a laboratory context and in the real world ([table 4](#)).

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 (100 Hz) 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

Table 4 Summary of the experimental protocol used for the validation of the algorithms in the laboratory and in the real world

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 Aeqora App Beacon	DynaPort MM+	Unsupervised real-world activities (including predetermined tasks)
Real world (7 days)	Mobile Phone with Aeqora App Beacon	DynaPort MM+	Unsupervised daily living

INDIP, INertial module with distance sensors and pressure insoles.

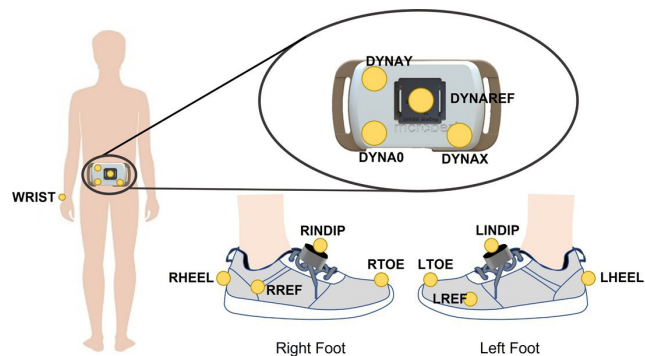


Figure 3 Illustration of the adopted marker set configuration. Markers were located on the right (RHEEL) and left (LHEEL) heels, toes (RTOE, LTOE) and on the INertial module with DIstance Sensors and Pressure insoles (INDIP) units located on the right and left foot (RINDIP, LINDIP). Two additional reference 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, DYNA0, DYNAX, DYNAREF).

instrumental errors, normally minimised to a few millimetres via ad hoc calibration procedures and filtering and smoothing techniques.²²

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*²³ will be used, which will establish the specific level of accuracy for each system. A graphical user interface for automated preprocessing 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 MM+ device (figure 3).

Tested devices

During this observation each participant will also be equipped with an additional multisensor system (INertial module with DIstance Sensors and Pressure insoles, INDIP)^{24–26} and with the Dynaport MM+. The INDIP system (figure 4) includes four inertial modules (one on

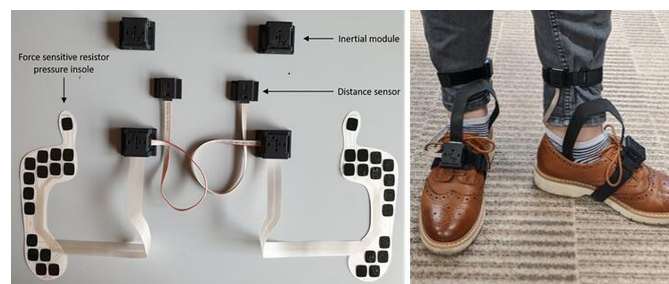


Figure 4 Different components of the INertial module with Distance Sensors and Pressure insoles system. The figure on the left shows the pressure insoles and the connectors that link them to the distance sensors and the inertial modules. The picture on the right shows how the same system is then attached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

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^{27,28} combined with a zero-velocity update,^{29,30} subsequently velocity and displacement will be calculated using a direct and reverse approach.^{24,31–33}

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.^{31,34,35} The final assembled system in its fully synchronised 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.³⁶

The lower back INDIP unit and the DynaPort MM+ will be rigidly attached to each other. The data from the SP (100 Hz), INDIP (IMU and insoles, 100 Hz, Distance sensor, 50 Hz) and DynaPort MM+ (100 Hz) 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

Strait walking: straight walking is the most common test of walking.^{37,38} The participant walks for a distance of 5 m from a standing start and will be repeated at three different walking speeds: preferred, fast and slow (figure 5A).

Timed Up and Go (TUG): the TUG is a widely used clinical assessment of a person's mobility.³⁹ 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 5B).

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 5C). Besides being a clinically validated test,⁴⁰ the main purpose of

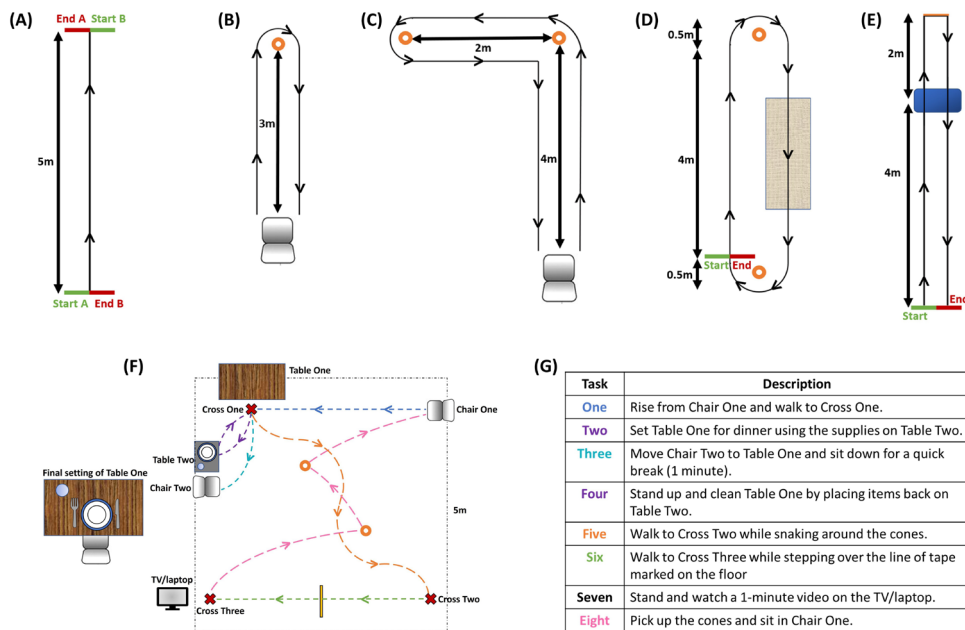


Figure 5 Diagrams of the selected tasks: (A) straight walking test, (B) timed up and go, (C) L-test, (D) surface test, (E) Hallway test, (F) schematic of the daily living activities, (G) description of the eight tasks performed during the daily living activities.

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 world:

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

Hallway test: the participant walks along a 6 m 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 5E).

Daily living activities

These lab-based tasks will be used to simulate daily activities expected in the real life, similar to previous studies.⁴¹ 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,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 simulated daily living activities, as well as handrails for the hallway test are permitted when needed.

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: (1) the core tracker, (2) the interface and (3) the server infrastructure collecting data across users. The core tracker, adapted from a library developed by the University of Sheffield,⁴² uses the mobile phone's internal sensors to compute the type of activity (eg, walking) and intensity (eg, cadence) to identify geo-located bouts of movement. It operates in the background and senses mobility features through a range of sensors (eg, step counters, activity recognition, accelerometer, gyroscope, etc) as well as from location services (Global Positioning System (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 these 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 OpenStreetMap⁴³ roads and paths, to remove GPS noise, the slope variation of each walk is computed on tiles, with a resolution of 5 m within the UK (using Ordnance Survey Terrain 5)⁴⁴ and 30 m in the other locations (using NASA's Shuttle Radar Topography Mission data,⁴⁵ 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 2 min); 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: 7 days monitoring

This observation will quantify the effects of device wearing time, hourly and daily fluctuations of DMOs, 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, if willingly). 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 7 days, without any specific instruction being provided, except for that of wearing the provided measurement tools.

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 7 day, free-living data collection, participants will complete two questionnaires to assess the acceptability of the device. The first is a 12-item questionnaire⁴⁶ 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,⁴⁷ 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 semistructured interview (see online supplemental file). For this qualitative part of the study, sampling will continue until saturation, that is, 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 Scale⁴⁸ 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 Questionnaire⁴⁹ (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

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

Variables	DMOs (units)	Definition	DMO attainable			
			DynaPort MM+	SP System	INDIP	Aeqora App
Walking bout (WB) <i>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).</i>	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) <i>Refers to the duration (time intervals) of strides/steps, calculated as the time in between two non-consecutive (alternate) initial contacts.</i>	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/min)	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	✓	✓	✓	

Continued

Table 5 Continued

Variables	DMOs (units)	Definition	DMO attainable		
	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	✓		✓

INDIP, INertial module with DIstance Sensors and Pressure insoles; SP, stereophotogrammetric.

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 semistructured interview with the aim of exploring their experiences of the data collection process. Assessors will be asked about the use of the device (eg, 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.

Data management

All data will be uploaded to a central platform 'e-Science Central' (e-SC)⁵⁰ which provides data processing and storage functionality in accordance with principles of reproducible research. The underlying infrastructure complies with the International Organization for Standardization (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 Administration⁵¹ for making them attributable, legible, contemporaneous, original and accurate. 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. 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 (eg, 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. The algorithms being developed and benchmarked will be used to process these files and extract and store the DMOs.

Data analysis plan

Verification of the device

Mean and SD readings of the accelerometer, gyroscope and magnetometer signals captured during static

Table 6 List of statistical analyses and performance metrics that will be used for the various digital mobility outcomes (DMOs)

DMO	Performance metric				Criterion validity				Plots						
	Sensitivity	Positive predictive value	Accuracy	Specificity	F1-score	Error (absolute and 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						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

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.

ICC, intraclass correlation coefficients; WB, walking bouts.

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 SD of nominal, measured and relevant errors for angular velocity values during the dynamic acquisitions.

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) 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 (eg, cadence, real-walking speed) will be summarised with descriptive statistics for the values obtained within walking bouts (mean and SD). ICC and its CI (95% CI) will be provided. In addition, the mean, minimum, maximum, SD, median, IQR and root mean square error of DMOs over all available walking bouts will be presented. Categorical variables (eg, laterality of initial contacts) will be summarised with frequency counts and percentages.

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 s⁵² and centred on each event detected by the reference system. The following performance metrics will then be calculated:

$$\begin{aligned} \text{Sensitivity} &= \frac{TP}{TP+FN} \\ \text{Positive Predicted Value} &= \frac{TP}{TP+FP} \\ \text{Accuracy} &= \frac{TP+TN}{TP+TN+FP+FN} \\ \text{Specificity} &= \frac{TN}{TN+FP} \\ \text{F1 score} &= 2 * \frac{\text{Positive Predicted Value} * \text{Sensitivity}}{\text{Positive Predicted Value} + \text{Sensitivity}} \end{aligned}$$

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:

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

$$\text{Absolute Error} = | \text{DMO estimated by IMU} - \text{DMO estimated by Reference System} |$$

The mean, SD 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 (eg, ICC (2,1)).

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

If participants do not participate in one assessment (eg, one of the tasks in the laboratory) or observation (eg, 2.5 hs), 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.⁵³

Participants' and assessors' experience

Participant and professional questionnaire data will be analysed using descriptive statistics. Interviews will be analysed using thematic analysis.⁵⁴ Transcripts will be deductively examined for the presence of themes related to the acceptability of the monitoring device to participants (ie, 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.

Author affiliations

¹INSIGNEO Institute for in silico Medicine, The University of Sheffield, Sheffield, UK

²Department of Mechanical Engineering, The University of Sheffield, Sheffield, UK

³Translational and Clinical Research Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne, UK

⁴Laboratory of Movement Analysis and Measurement, Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland

⁵Robert Bosch Gesellschaft für Medizinische Forschung, Stuttgart, Germany

⁶Department of Biomedical Sciences, University of Sassari, Sassari, Sardegna, Italy

⁷The Newcastle Upon Tyne Hospitals NHS Foundation Trust, Newcastle Upon Tyne, UK

⁸Center for the Study of Movement, Cognition and Mobility, Neurological Institute, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel

⁹ISGlobal, Barcelona, Spain

¹⁰Universitat Pompeu Fabra (UPF), Barcelona, Spain

¹¹CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

¹²IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain

¹³Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Torino, Italy

¹⁴PolitoBIOMed Lab – Biomedical Engineering Lab, Politecnico di Torino, Torino, Italy

¹⁵Insight Centre for Data Analytics, O'Brien Science Centre, University College Dublin, Dublin, Ireland

¹⁶UCD School of Public Health, Physiotherapy and Sports Science, University College Dublin, Dublin, Ireland

¹⁷Department of Electrical, Electronic and Information Engineering «Guglielmo Marconi», University of Bologna, Bologna, Italy

¹⁸Health Sciences and Technologies—Interdepartmental Center for Industrial Research (CIRI-SDV), University of Bologna, Bologna, Italy

¹⁹Department of Sport, Exercise and Rehabilitation, Northumbria University Newcastle, Newcastle upon Tyne, UK

²⁰Department of Computer Science, The University of Sheffield, Sheffield, UK

²¹Machine Learning and Data Analytics Lab, Department of Artificial Intelligence in Biomedical Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

²²McRoberts BV, Den Haag, Zuid-Holland, Netherlands

²³Department of Neurology, University Medical Center Schleswig-Holstein Campus Kiel, Kiel, Germany

²⁴Department of Physical Therapy, Sackler Faculty of Medicine & Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

²⁵Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, Trondheim, Norway

²⁶School of Computing, Newcastle University, Newcastle upon Tyne, UK

²⁷Novartis Institutes of Biomedical Research, Novartis Pharma AG, Basel, Switzerland

²⁸Department of Neuroscience and Sheffield NIHR Translational Neuroscience BRC, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

²⁹Digital Health R&D, AstraZeneca Sweden, Sodertalje, Sweden

Twitter Emily Hume @emilyhume and Felix Kluge @kluger_felix

Contributors 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 multidisciplinary 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, Ake, KS, SB, FC, LRo. Design and deployment of experimental tools and protocols: KS, LS, CM, TB, EB, FC, NI, BS, LVG, IV, EH, NC, FS, AC, MC, SB, LRo, AY, CK, M-EMA, LS, CB, WM, CH, EW, JE, EG, MB, JLH, MN, KT, BV, JMH, Ake, DS, AM, VL. Data collection: LA, TB, PB, MB, EB, FC, EG, CH, EH, WM, CB, LPI, LS, KS, BS, LVG, IV, AY, EW. Data analysis and algorithm development: AP-I, SDD, M-EMA, HS, AS, BE, FK, AKü, MU, AC, JGA, SK, LC, LPa, LRe, HH, FC, EG. Draft revisions: LRo, KA, BE, JMH, BV, HS, AM, JLH, PB, FK, BC, Ake, LC, LPI, EB, KS, SK, LPa, HH, WM, NC, BS, AP-I, IN, VL, SDD. Study design and coordination: CM, LRo, AM, SDD, AC. All authors have read and approved the final manuscript.

Funding 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, AY and LRo 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. 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.

Competing interests FC is CEO and shareholder of Aeqora Ltd, of which VL is director and shareholder. HS is an employee of AstraZeneca. BE is co-founder and owns shares of Portables HealthCare Technologies GmbH. McRoberts is the manufacturer of the DynaPort. MN, JE, and LPI are employees of McRoberts. AM is an employee of Novartis and holds stock in Novartis. LS and CB are consultants of Philipps Healthcare, Bosch Healthcare, Eli Lilly, Gait-up. JMH 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. LPa and LC are co-founders and own shares of mHealth Technologies.

Patient consent for publication Consent obtained directly from patient(s).

Provenance and peer review Not commissioned; externally peer reviewed.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Claudia Mazzà <http://orcid.org/0000-0002-5215-1746>

Ellen Buckley <http://orcid.org/0000-0002-0968-6286>

Anne-Elie Carsin <http://orcid.org/0000-0003-3981-977X>

Clint Hansen <http://orcid.org/0000-0003-4813-3868>

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.
- 3 Perera S, Patel KV, Rosano C, *et al*. Gait speed predicts incident disability: a pooled analysis. *J Gerontol A Biol Sci Med Sci* 2016;71:63–71.
- 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 Biomark* 2020;4:13–27.
- 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.
- 6 Bonci T, Keogh A, Del Din S, *et al*. An objective methodology for the selection of a device for continuous mobility assessment. *Sensors* 2020;20:1–16.
- 7 Gao Y, Li H, Luo Y. An empirical study of wearable technology acceptance in healthcare. *Ind Manag Data Syst* 2015;115:1704–23.
- 8 Jia Y, Wang W, Wen D, *et al*. Perceived user preferences and usability evaluation of mainstream wearable devices for health monitoring. *PeerJ* 2018;6:1–17.
- 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:e173.
- 10 Viceconti M, Hernandez Penna S, Dartee W, *et al*. Toward a regulatory qualification of real-world mobility performance biomarkers in Parkinson’s patients using digital mobility outcomes. *Sensors* 2020;20. doi:10.3390/s20205920. [Epub ahead of print: 20 Oct 2020].
- 11 Mobilise-D website, 2020. Available: <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.
- 13 Letter of support for Mobilise-D digital mobility outcomes as monitoring biomarkers. Available: 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.
- 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. *Sens Actuators A Phys* 1994;42:446–9.
- 19 Stancin S, Tomazic S. Time- and computation-efficient calibration of MEMS 3D accelerometers and gyroscopes. *Sensors* 2014;14:14885–915.

- 20 Kluge F, Del Din S, Cereatti A, *et al.* Consensus based framework for digital mobility monitoring. *PLoS One* 2021;16:12.
- 21 COSMIN. Available: <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.
- 23 Di Marco R, Rossi S, Castelli E. 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.
- 24 Trojaniello D, Cereatti A, Della CU. *Gait direction of progression changes using Shank worn MIMUs*. GBN Conference, Pavia: In, 2014.
- 25 Salis F, Bertuletti S, Caruso M. Multi-sensor integration and data fusion for enhancing gait assessment in and out of the laboratory. *Gait Posture* 2019;74:34.
- 26 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.
- 27 Caruso M, Sabatini AM, Laidig D, *et al.* 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.
- 28 Caruso M, Sabatini AM, Knaflitz M. 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.
- 29 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.
- 30 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.
- 31 Trojaniello D, Cereatti A, Pelosin E. *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.
- 32 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;9:9.
- 33 Trojaniello D, Cereatti A, Ravaschio A. 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.
- 34 Betoli M, Cereatti A, Trojaniello D. Estimation of spatio-temporal parameters of gait from magneto-inertial measurement units: multicenter validation among Parkinson, mildly cognitively impaired and healthy older adults. *Biomed Eng Online* 2018;17:1–4.
- 35 Rossaningo R, Caruso M, Salis F. An optimal procedure for Stride length estimation using Foot-Mounted Magneto-Inertial measurement units. In: *2021 IEEE International Symposium on medical measurements and applications (MeMeA)*, 2021: 1–6.
- 36 Salis F, Bertuletti S, Caruso M. A novel multi-sensor system for gait assessment in real-world conditions: preliminary results. *Proceedings of GNB 2020*.
- 37 Zijlstra W, Hof AL. Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture* 2003;18:1–10.
- 38 Ngueleu AM, Blanchette AK, Bouyer L, *et al.* Design and accuracy of an instrumented insole using pressure sensors for step count. *Sensors* 2019;19:12–14.
- 39 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* 2019;27:230–3.
- 40 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.
- 41 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.
- 42 Ciravegna F, Gao J, Ireson N. 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.
- 43 Open street MAP. Available: <https://www.openstreetmap.org> [Accessed 18 Feb 2021].
- 44 Ordnance survey terrain 5. Available: <https://www.ordnancesurvey.co.uk/business-government/products/terrain-5> [Accessed 18 Feb 2021].
- 45 U.S. Releases enhanced shuttle land elevation data. Available: <https://www2.jpl.nasa.gov/srtm/> [Accessed 18 Feb 2021].
- 46 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.
- 47 Knight JF, Baber C, Schwirtz A. The comfort assessment of wearable computers. *Proc - Int Symp Wearable Comput ISWC* 2002:65–72.
- 48 Brooke J. SUS: a "quick and dirty" usability". *Usability Eval Ind* 1995;11:189.
- 49 Lewis JR. IBM computer usability satisfaction questionnaires: psychometric evaluation and Instructions for use. *Int J Hum Comput Interact* 1995;7:57–78.
- 50 Hiden H, Woodman S, Watson P, *et al.* Developing cloud applications using the e-Science central platform. *Philos Trans A Math Phys Eng Sci* 2013;371:20120085.
- 51 Data Integrity and Compliance With CGMP. Guidance for industry. Available: <https://www.fda.gov/files/drugs/published/Data-Integrity-and-Compliance-With-Current-Good-Manufacturing-Practice-Guidance-for-Industry.pdf> [Accessed 18 Feb 2021].
- 52 Tietsch M, Muaremi A, Clay I, *et al.* Robust step detection from different Waist-Worn sensor positions: implications for clinical studies. *Digit Biomark* 2020;4:50–8.
- 53 Kang H. The prevention and handling of the missing data. *Korean J Anesthesiol* 2013;64:402.
- 54 Chiari L, Della Croce U, Leardini A, *et al.* Human movement analysis using stereophotogrammetry. Part 2: instrumental errors. *Gait Posture* 2005;21:197–211.
- 55 Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3:77–101.
- 56 Postuma RB, Berg D, Stern M, *et al.* Mds clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2015;30:1591–601.
- 57 Jette AM, Haley SM, Coster WJ, *et al.* Late life function and disability instrument: I. development and evaluation of the disability component. *J Gerontol A Biol Sci Med Sci* 2002;57:M209–16.
- 58 Haley SM, Jette AM, Coster WJ, *et al.* Late life function and disability instrument: II. development and evaluation of the function component. *J Gerontol A Biol Sci Med Sci* 2002;57:M217–22.
- 59 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.
- 60 Kurtzke JF. Rating neurologic impairment in multiple sclerosis. *Neurology* 1983;33:1444 LP–1444.
- 61 Gupta N, Pinto LM, Morogan A. The COPD assessment test: a systematic review. *Eur Respir J* 2014;44:873 LP–84.
- 62 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.