The University Of Sheffield.

Measuring gait in the real world: challenges and solution for a laboratory based technical validation

Kirsty Scott, M.Sc.

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield

Faculty of Engineering

Department of Mechanical Engineering

February 2023

SUMMARY

Wearable motion sensors are leading the transition from the traditional laboratory or clinical based assessment, to quantifying mobility in free living environments. However, the validity of these so-called digital mobility assessment tools is limited, especially within patient populations. A primary reason for this it that there is the complexity in creating a validation protocol that encompasses all aspects of real-world while maintaining the integrity of validation measure. This is particularly crucial for the development of digital mobility assessments as a clinically valid tool in patient cohorts with slow walking speeds and abnormalities in gait.

This PhD will contribute to the development of a study for the technical validation of a digital mobility assessment device and subsequent quantification of the reliability and consistency of the calculated digital mobility outcomes (DMOs). Within this context, the specific aim of this thesis is to define, validate and deploy, as part of the above technical validation, an experimental protocol and the relevant algorithmic and laboratory tools needed to enable a robust multicentric collection of gait data from a wide range of disease groups.

Prior to the development of the validation study, a systematic review was completed to assess the current state of the art in validation protocols, with aim of highlighting common concepts and limitations within the literature to aid the development of a protocol framework. Which highlighted a need for standardisation of a gold standard solution and the use of a multi-stage and multi-context protocol to ensure a data collection that encompassed all aspects of real-world gait, while promoting a statistically strong validation.

In light of these requirements, standardisation of a gold standard solution for the validation of a wearable sensor within a lab environment and a multi-stage and multi-context protocol were developed and validated for creating this protocol framework and subsequently the successful validation of a wearable sensor solution for estimating real-world gait.

ACKNOWLEDGEMENTS

First and foremost, I am extremely grateful to my supervisor Prof. Claudia Mazzà for her invaluable advice, continuous support, and patience during my PhD. You have always been open to sharing your knowledge and general insights, which this achievement would not have been possible without. Thank for your words of wisdom and confidence boosts during the low points and being there to celebrate the highs. My thanks to Dr Jen Rowson and Prof Andrea Cereatti for their continued support in my PhD development and all their guidance in getting me over the finish line.

I am especially thankful to the IMSB group at the University of Sheffield. The ability to work with such a multidisciplinary group of researchers has opened my eyes to the wealth of possibilities research offers in improving healthcare. Thank you for giving me this opportunity, for all the friendships I have made and for the much-needed social events you have all organised. A special thank you goes to Dr Tecla Bonci for being a strong mentor, great colleague, and wonderful friend from day one. I am so grateful to have had you to laugh, moan and celebrate with throughout this PhD.

Thank you to the Mobilise-D consortium for allowing me this opportunity, with a special thank you to Grünenthal GmbH for financially supporting me through my doctoral studies and their continued interest in my progress. Thank you to all the members of WP2 for the amazing support and teamwork during the technical validation study.

From a personal note, I would like to thank my husband Nick for his support in my career and the confidence he has given me to follow my own path. Your continued words of reassurance, hugs and home cooking have been the boosts I have needed. I am so grateful to have you as a partner.

Thank you to my brother Euan and my dear friend Natalie for all their love and support throughout this chapter.

Special thank you to my parents, Cathy and Tommy, who have been there from the start! The love and guidance you have given me is insurmountable. Thank you for always being there when I need you and giving me the opportunities that led me here.

Finally, I would like to thank all the patients I have worked with in my clinical work and those who have supported this research with their participation. Thank you for giving me the drive to see me through this PhD and for your honesty in sharing your healthcare experiences. Without your support this would not have been possible.

PUBLICATIONS

Scientific Publications

Scott K, Bonci T, Alcock L, Buckley E, Hansen C, Gazit E, Schwickert L, Cereatti A, Mazzà C, on behalf of the Mobilise-D Consortium. A Quality Control Check to Ensure Comparability of Stereophotogrammetric Data between Sessions and Systems. Sensors. 2021; 21(24):8223. doi: 10.3390/s21248223

Scott, K., Bonci, T., Salis, F.et al. Design and validation of a multi-task, multi-context protocol for real-world gait simulation. J NeuroEngineering Rehabil 19, 141 (2022). https://doi.org/10.1186/s12984-022-01116-1

Polhemus, A. M., Bergquist, R., de Basea, M. B., Brittain, G., Buttery, S. C., Chynkiamis, N., ... & Frei, A. (2020). Walking-related digital mobility outcomes as clinical trial endpoint measures: protocol for a scoping review. BMJ open, 10(7), e038704.

Bonci, T., Keogh, A., Del Din, S., Scott, K., Mazzà, C., & Mobilise-D Consortium. (2020). An objective methodology for the selection of a device for continuous mobility assessment. Sensors, 20(22), 6509.

Polhemus, A., Ortiz, L. D., Brittain, G., Chynkiamis, N., Salis, F., Gaßner, H., ... & Puhan, M. (2021). Walking on common ground: a cross-disciplinary scoping review on the clinical utility of digital mobility outcomes. NPJ digital medicine, 4(1), 1-14.

Mazzà, C., Alcock, L., Aminian, K., Becker, C., Bertuletti, S., Bonci, T., ... & Rochester, L. (2021). Technical validation of real-world monitoring of gait: a multicentric observational study. BMJ open, 11(12), e050785.

Bonci, T., Salis, F., Scott, K., Alcock, L., Becker, C., Bertuletti, S., ... & Mazzà, C. (2022). An algorithm for accurate marker-based gait event detection in healthy and pathological populations during complex motor tasks. Frontiers in bioengineering and biotechnology, 10.

Micó-Amigo, M. E., Bonci, T., Paraschiv-Ionescu, A., Ullrich, M., Kirk, C., Soltani, A., ... & Del Din, S. (2022). Assessing real-world gait with digital technology? Validation, insights and recommendations from the Mobilise-D consortium. (Under review).

CONTENTS

NOMENCLATURE

DMO	Digital Mobility Outcome
IC	Initial Contact
FC	Final Contact
WB	Walking Bout
WS	Walking Speed
TUG	Timed Up and Go
6MWT	Six Minute Walking Test
IMU	Inertial Measurement Unit
MIMU	Magneto Inertial Measurement Unit
GRF	Ground Reaction Force
SP	Stereo Photogrammetric
GCS	Global Coordinate System
LCS	Local Coordinate System
GUI	Graphical User Interface
HA	Healthy Adults
MS	Multiple Sclerosis
PD	Parkinson's Disease
PFF	Proximal Femoral Fracture
COPD	Chronic Obstructive Pulmonary Disease
CHF	Congestive Heart Failure

1 INTRODUCTION

1.1 Importance of Mobility

According to the World Health Organisation definition, mobility is “the activity of moving by changing body position or location or by transferring from one place to another, by carrying, moving or manipulating objects, by walking, running or climbing and by using various forms of transportation” (World Health, 2001). As such, a decreased level of mobility is highly linked to decline in a person’s quality of life, regarding both their physical health and general well-being (Perera et al., 2015; Studenski et al., 2011). A decrease in mobility is a common symptom of aging (Gill et al., 2006) but can also present as a clinical symptom in certain diseases that affect the necessary organs or neurological functions required in mobility (Sue Lord et al., 2013). In this instance, a change in mobility can be one of the first signs of disease even prior to reported symptoms, and as such, has the potential to be a key indicator in the early diagnosis of certain pathologies (Buckley et al., 2019; Rochester et al., 2014), making it a powerful measure to consider in healthcare.

As defined in Del Din et al., (2016), a person’s level of mobility can be measured in an accumulative sense, i.e., the range of motor tasks and movements performed in the day-to-day (macro analysis), or it can be broken down to focus on a specific function, to gain a more comprehensive understanding of the strategies used in performing certain tasks such as walking, also referred to as gait (micro analysis). Both approaches offer the opportunity to monitor longitudinal changes at subject level and observe characteristics at group level which may highlight areas of focus for disease specific treatments. However, a micro-analysis that includes a specific function of mobility has shown a higher level of sensitivity to subtle changes or deterioration that would not likely be caught as early when considering mobility as a whole (Silvia Del Din, Aodhan Hickey, et al., 2016). As such, analysis centring on functions of mobility, in particular gait, has become a major focus in research, healthcare, and the pharmaceutical industry (Rochester et al., 2020).

1.2 Gait: Definitions and Background

Gait is the forward movement of a person by use of the lower limbs. Traditionally, gait is investigated by breaking it down into cycles, which include a sequence of events that start with the foots initial contact (IC) with the ground to its subsequent IC with the ground (Figure 1.1).

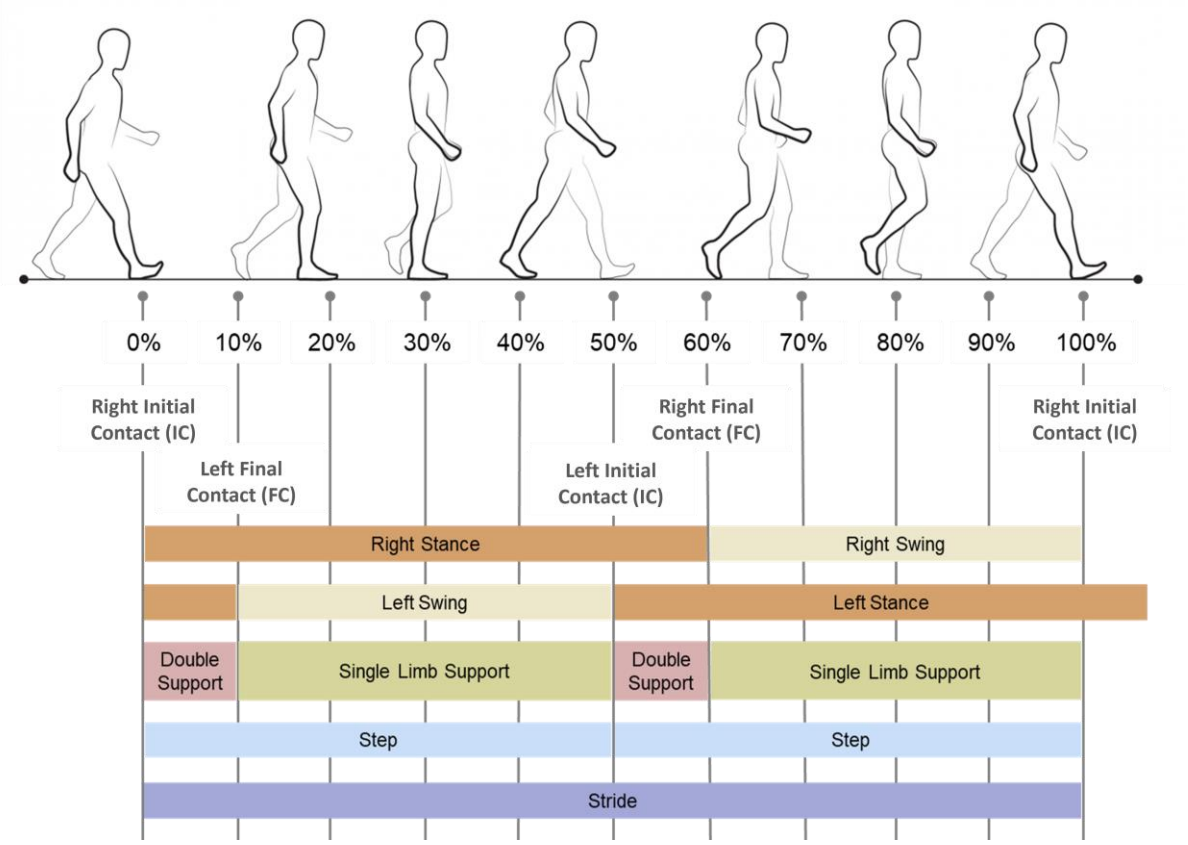


Figure 1.1. The gait cycle, with the right leg dominant and depiction of the events and phases completed (adapted from (APDM, 2021)).

The gait cycle is split into two phases: stance and swing. The stance phase (60% of the gait cycle) begins with the IC and ends with the final contact (FC) of the same foot. During this phase, the foot is weight bearing and primarily acts as a single support for the body. However, during the start and end of stance phase there is a double support due to the FC and IC of the contralateral leg. The swing phase (40% of the gait cycle) is the period between the FC and subsequent IC of the same foot. This point in the cycle is crucial in gaining the forward momentum required in gait.

Using the gait events and phases described above, it is possible to extract temporal, spatial and spatiotemporal parameters which can be useful measures in characterising a person's efficiency in walking. These include:

- ***Stance duration***: time between IC and FC of the same leg
- ***Swing duration***: time between the FC and IC of the same leg
- ***Step duration***: time between two consecutive IC events of the different legs.
- ***Stride duration***: time between two consecutive IC events of the same leg.
- ***Step Length***: distance between two consecutive IC events of the different legs.
- ***Stride Length***: distance between two consecutive IC event of the same leg.
- ***Cadence***: the number of steps per unit time
- ***Walking Speed***: the ratio of distance covered against time taken.

As described above, gait is considered as a cyclic movement with distinguishable events that become relatively automatic motor functions in predictable environments. However, compensation to this cycle are likely when the act of walking becomes more complex due to changes in the environment (i.e., walking surface, changes in weather or crowded/confined spaces), increased complexity of the task (i.e., dual tasks, turns, inclines/declines) or as mitigation to the effects of aging or pathology (Holtzer et al., 2006). To evaluate alterations made to combat any of the above factors, it is crucial to consider gait as a set of consecutive cycles completed during a continuous period of walking, referred to in this thesis as a walking bout (WB). The assessment of spatiotemporal parameters over a WB, offers the opportunity to consider a holistic assessment of a person's gait over multiple steps, that accounts for variability, stability, rhythm, pace, and asymmetry, all of which are domains that have been shown to be key characteristics of gait change linked to pathologies and clinical endpoints such as falls risk (Costa et al., 2022; S. Lord et al., 2013). In addition, this approach can allow comparison between multiple WBs within or between subjects, aiding assessment of changes in walking over time (whether that be to assess daily fluctuations or longitudinal change), in different environments (e.g., indoor or outdoor, changes in weather or temperature), in tasks with varying complexity (e.g., straight walking, inclines or postural transitions) and possibly the effect of an intervention (e.g., pharmacological, assistive devices or rehabilitation).

When assessing gait across a WB, it is crucial to ensure that there is a clear definition of what constitutes as a WB and how WBs are characterised in the context of the assessment (i.e., grouping between indoor and outdoor WBs or grouped based on varying difficulty of task). Until recently, the level of heterogeneity within the research regarding these definitions has limited the comparability of data across cohorts and studies and as such is a widely discussed as a limitation in the literature(S. Del Din et al., 2016; Espay et al., 2019). In response, Kluge et al., (2021) gathered a consensus from 162 academic, clinical and industrial experts in the field of gait analysis, to reach an agreement on the classification of various descriptive phrases

and definition of mobility outcomes, related to mobility monitoring. The result included agreement on a range of terms and minimum requirements for the calculation of certain gait related outcomes, including the definition of a WB (a minimum of two gait cycles), which will be adopted throughout the analysis presented in this thesis (Figure 1.2).

Term	Aspect	Definition
Walking	Physiological	Human **walking** is a modality of locomotion and is defined as initiating and maintaining a forward displacement of the centre of mass in an intended direction involving the use of the two legs, which provide both support and propulsion. The feet are repetitively and reciprocally lifted and set down whereby at least one foot is in contact with the ground at all times **Walking** with walking aids is included in this definition. A **step** is the interval between the initial contacts of the ipsi- and contralateral foot and corresponds to the forward displacement of the foot together with a forward displacement of the trunk . A **stride** is the interval between two successive initial contacts of the same foot. As such, a **stride** is equivalent to the gait cycle and every stride contains two **steps**
	Relation to walking bouts	**Walking** is made up of walking bouts and is equivalent to taking steps/stepping forward (thus stepping in place does not constitute walking) and is defined as starting from initial contact for the initial step until ending with full floor contact of the foot making the last step
Purposeful	Characteristics	**Purposeful walking** includes an intentional component of the movement (e.g., getting to the bathroom, catching the bus, going to the grocery store, going for a walk in the park, etc.). **Purposeful walking** may constitute certain characteristics (e.g., more constant walking velocity, lower variability of gait characteristics, straighter direction of locomotion than non-purposeful walking, specific context, etc.). Those gait characteristics are quantified based on discrete walking bouts.
Real-world	Characteristics	**Real-world** relates to the context in which walking takes place—that is free-living, unsupervised, uncontrolled and non-standardised. As such, it is unscripted as there are no instructions to the subject who does not need to interact with the wearable device(s). **Real-world** actions occur in non-simulated everyday situations in unconstrained environments with minimal consciousness of being tested. It is equivalent to actions at home or in the community over continuous periods of time Synonymous terms are (environment of) **daily living**, or relating to **daily-life**. **Home environment** is used synonymously to **real-world** and **daily-life** without a separation of indoor and outdoor environment **Real-world** is distinct from laboratory-based , supervised (= fully controlled and observed), and semi-controlled (walking'freely' but with supervision) tests. It also is different from scripted/instructed walking, which can take place in the home or lab (such as walking tests like the 4x10m test, 6-minute walk test (6MWT) and timed up and go (TUG)).
	Clinical environment	Free walking in hospitals is part of the **real-world** definition, but standardized supervised tests in a hospital are not. This excludes instructed actions, e.g., by medical professionals.
	Standardized measurement	Home-based tests, which are semi-standardized measurements performed in the home environment in a controlled or semi-controlled environment (such as short walk tests), are thus not regarded as being part of **real-world**. Home-based tests can nevertheless be an alternative to clinical tests and might be easier to conduct operationally and analyse than continuous monitoring (assuming standardized instructions).
Walking bout	Characteristics	A **walking bout (WB)** is a walking sequence containing at least two consecutive strides of both feet (e.g., *R-L-R-L-R-L* or *L-R-L-R-L-R*). Start and end of a **walking bout** are determined by a resting period or any other activity (non-walking period). The initial step of a **WB** follows a non-walking period and the final step precedes the next non-walking period.
Walking speed	Physical definition	**Walking speed (WS)** is the distance covered by the whole body within a certain time interval / per unit time of walking. It is measured in meters per second and is the magnitude of the velocity vector (velocity includes direction and magnitude of walking)
	Granularity	**Walking speed** can be estimated at different granularities: • Instantaneous WS varies from one instant to another during the walking cycle • Step-wise WS is the ratio between step distance (length) and step time • Stride-wise WS • Averaged over WBs • Averaged over other time intervals (hourly, daily, weekly) based on multiple WBs The granularity by which the WS is assessed should be related to clinical parameters for each population separately.
	Relation to walking bouts	**Walking speed** will be assessed with regard to walking bouts. Thus, the minimal length of one walking bout required to assess average walking speed is based on a sequence of 2 consecutive strides (e.g., *R-L-R-L-R-L* or *L-R-L-R-L-R*).
Turning	Characteristics	The process of **turning** consists of decelerating the forward motion, rotating the body as a whole, and stepping out toward the new direction . Thus, **turning** includes a change of walking direction and change in angular orientation including a rotational movement of the body around the longitudinal axis. **Turning**, curvilinear walking, and straight walking involve different neuromotor strategies and need to be discriminated.

Figure 1.2 Agreed definitions of terms related to the assessment of gait (Kluge et al., 2021).

1.3 Common Approaches to Measuring Gait

Based on the wealth of health-related information available from assessing gait, it is of no surprise that measurements of mobility, that incorporate gait specific metrics, are commonly used across healthcare and research. However, these measures come with their limitations. The primary being that many are carried out in the form of patient reported outcomes, such as the International Physical Activity Questionnaire (IPAQ) (Booth, 2000) or the Late-life Disability Index (LLFDI) (Jette et al., 2002). Although these measures may give a general understanding to the person's ability to move and perceived challenges during functions like gait, they can be highly sensitive to errors and have shown poor levels of repeatability (Tomioka et al., 2011).

An alternative approach is the addition of a short walking test, such as the Timed Up and Go (TUG) (Podsiadlo & Richardson, 1991), the 6 Minute Walking Test (6MWT) (Crapo et al., 2002) or a test integrated within disease specific clinical assessments such as the Unified Parkinson's Disease Rating Scale (MDS-UPDRS) (Goetz et al., 2008) or the Expanded Disability Status Scale (EDSS) (Kurtzke, 1983).This form of test aims to capture one of the qualifiers of mobility, mobility capacity (i.e., the highest level of mobility that an individual may achieve in each standardised environment), with the other qualifier, mobility performance (i.e., what people do in their current environment), only being assessable by monitoring an individual in real-world conditions.

In these capacity-based tests, the patient is required to perform a structured task under the supervision of a clinician. For instance, in the 6MWT participants are asked to walk as far as they can up and down walkway in 6 minutes (Crapo et al., 2002). Traditionally, these tests measure the spatiotemporal elements of walking via a stopwatch, measured distance, or manual step count (either by eye, a handheld clicker or 2D video). However, in more recent years, researchers have started instrumenting these assessments using more sophisticated measuring systems, such as optoelectronic systems that track movement using passive markers attached to the body and walkways or treadmills with integrated force platforms, to allow the quantification of a larger variety of gait parameters (Givon et al., 2009; Rucco et al., 2017; Sosnoff et al., 2011). Although this form of instrumentation offers a more detailed, quantitative measure of gait with a high level of accuracy, these systems are not readily available within clinical settings, due to cost and the expertise needed to run and process the outputted data. In addition, these systems can be restrictive in their use, only allowing analysis of periods of straight walking, which may not induce common strategies used by the patient in the more complex aspects of the test (such as turning in the 6MWT and TUG) that may be of pertinence to the progression or early signs of a pathology.

Although the reliability of using the traditional tools in these tests, over repeated measures and between operators, has previously been shown to be minimal to the clinical meaningfulness of the measure (Morris et al., 2001), this testing format is limited to a small amount of gait related outputs, subsequently limiting its sensitivity to subtle changes in gait function when compared to the more sophisticated methods of instrumentation, even with restriction in the analysis to the periods of straight walking (Martin et al., 2006).

Regardless of its limitations, this form of testing has proven to be an effective way of measuring mobility capacity, but could benefit from more sophisticated measuring tools, that are readily available and can continuously monitor throughout the task, to improve the efficiency of the assessment's outcome (i.e., stratification of disease).

1.4 Measuring Gait with Wearable Sensors

Wearable motion sensors, such as an inertial measurement unit (IMU) or magneto inertial measurement unit (MIMU) are relatively cheap, easy to use and are not restrictive in application or environment, in comparison to the traditional laboratory-based methods for measuring gait ,mentioned in the section above. As such, wearable sensors are a promising tool for continuously assessing gait, both in the context of capacity tests but also as a means of monitoring mobility performance in the real-world.

1.4.1 Traditional Inertial Sensor Components

An IMU is a device that integrates a 3D accelerometer and 3D gyroscope with 3 orthogonal sensitive axes. A MIMU is an extended version that encompasses magnetometers, measuring the earth's magnetic field. The addition of a magnetometer can allow for insight into the sensor's orientation (and consequently the body segment it is attached to) in respect to a fixed global frame. For example, if the system is attempting to quantify and characterise gait, the magnetometer can add depth to the analyses by determining the direction the person is walking. The principles of a magnetometer are based on the Lorentz force, which describes the force generated by charged electrons in a current when placed in a magnetic field (Lowe & ÓLaighin, 2014).

Accelerometers: sense linear acceleration along one or several axes and are composed of a proof mass attached to a mechanical suspension system, with respect to a reference frame. When the accelerometer experiences acceleration, inertial force will cause the mass to deflect according to Newton's Second Law. Acceleration is measured electrically by calculating the displacement of the mass with respect to the reference frame (Yang & Hsu, 2010). There are two categories of accelerometers: piezoelectric (Figure 1.3) and capacitive

based, with both previously being utilised for assessing human movement (Lowe & ÓLaighin, 2014).

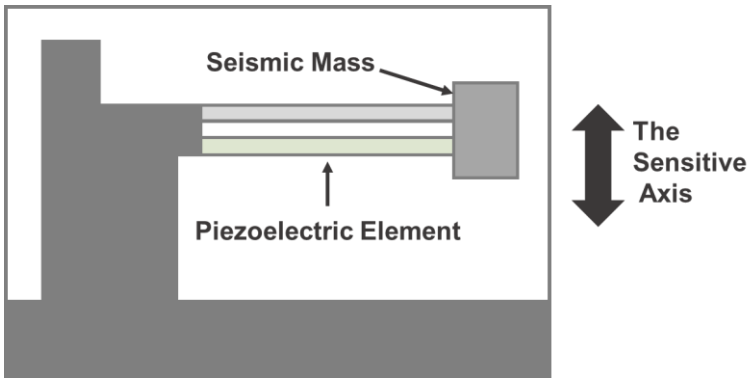


Figure 1.3. Piezoelectric accelerometer

Gyroscopes: sense angular motion about one or several axes directly proportional to the angular rate applied. The ability for a gyroscope to provide this information is reliant on the Coriolis effect, which describes the deflection of a moving object when viewed from a moving reference point (Lowe & ÓLaighin, 2014). The Coriolis force is a vector quantity that is perpendicular to the velocity and rotation of the object and has a magnitude proportional to the angular velocity according to:

$$\boldsymbol{F_c = -2m(\omega \times v)}$$

Where F_c is the Coriolis force, ω is the angular velocity, m is the mass of the moving object and v is the linear velocity of the movement. Gyroscopes are produced in various forms with the vibrating fork (Figure 1.4) found to be the most used in assessing human movement (Lowe & ÓLaighin, 2014).

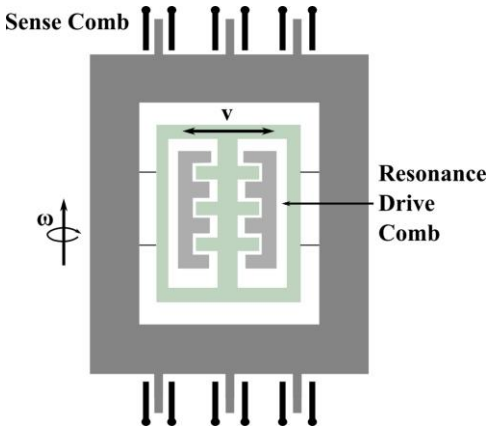


Figure 1.4. Vibrating fork gyroscope (adapted from (Lowe & ÓLaighin, 2014))

An additional sensor sometimes found in IMUs and MIMUs is a barometer. By transforming the change in pressure into a change in height, the use of a barometer can allow for the

detection of tasks such as ascending or descending of stairs. The most common form of barometer uses a diaphragm that deflects based on the difference between atmospheric pressure and reference pressure (Lowe & ÓLaighin, 2014).

1.4.2 Calculation of Gait Parameters

The calculation of gait parameters through systems like IMUs are commonly referred to as digital mobility outcomes (DMOs) and can include all the above mentioned spatial temporal parameters mentioned in Section 1.2 as well as specific outcomes for the gait domains of variability, stability, rhythm, pace, and asymmetry.

When recording gait, the 3D signals from the IMU or MIMUs components produce features that are linked to the specific events (such as ICs and FCs) found within the gait cycle. By identifying these features, the temporal events and phases in gait can be estimated, as shown in the example in Figure 1.5.

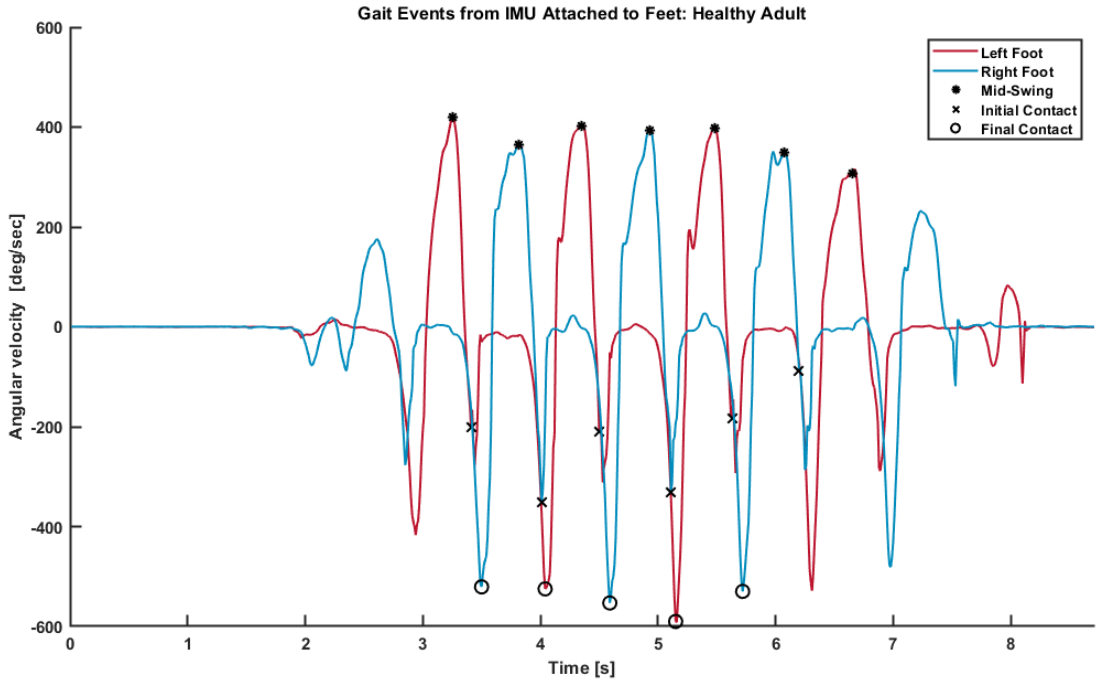


Figure 1.5Vertical angular velocity recorded by an IMU placed on the left (red) and right (blue) foot of a healthy participant while walking. FC (circle) , IC (cross) and mid-swing (star) events identified, based on the methodology of (Salarian et al., 2004).

To estimate the spatial parameters of gait, double integration of the IMU and MIMU components can be calculated. As example, gravity can be subtracted from the accelerations based on the device's rotation, and then integration of the residual acceleration can be

applied to gain the velocities. Finally, this integration is completed once more to gain the displacement and subsequently used to calculate the spatial features of gait (e.g., step length). This approach works well for IMUs or MIMUs attached directly to the foot, with additional modelling approaches used to transfer this estimation for sensors attached higher up the body (Zijlstra & Hof, 2003).

With advancements in techniques to extract features from these signals, such as personalisation to threshold crossings (Figueiredo et al., 2018), optimisation to the algorithms described briefly above (Zijlstra, 2004) and the addition of methods such as machine learning (Romijnders et al., 2022), there is scope for this form of measuring tool to meet a similar, if not the same, level of accuracy to the typical systems used in quantifying DMOs, while having the added bonus of being a more versatile and efficient tool for universal application.

1.4.3 Current Limitations of IMU-Based Methods

Based on the potential IMUs have in enriching analysis of the already clinically validated mobility capacity tests, there has inevitably been a large adoption of their use and validation within clinical-based settings (Khalaf et al., 2022). However, the correct application of these sensors does not go without its challenges. First, validation of the outputs and corresponding algorithms is required for each new pathological cohort of interest (Maetzler et al., 2016). The reason being that IMU components are sensitive to subtle changes in movement, meaning that the addition of disease-specific characteristics of gait (e.g., freezing of gait, shuffling or ataxia) can alter the shape of the signal and induce a larger signal to noise ratio , which may cause difficulties in extracting the gait specific features from the signal (as shown in the example of a recording of a person with multiple sclerosis in Figure 1.6 when compared to the previous example in Figure 1.5).

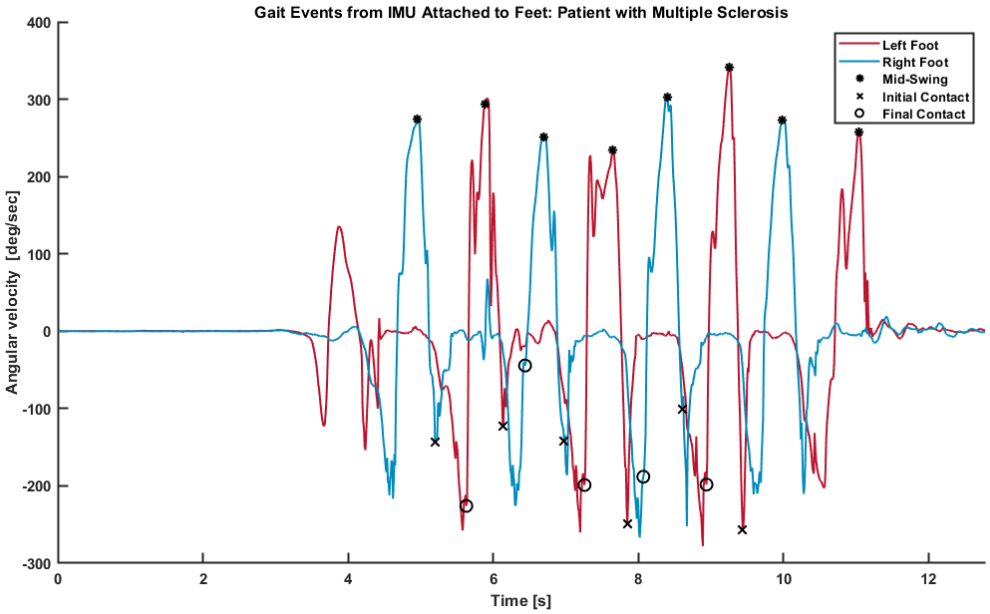


Figure 1.6 Vertical angular velocity recorded by an IMU placed on the left (red) and right (blue) foot of a participant with multiple sclerosis while walking. Gait events identified using same method as Figure 1.5 (Salarian et al., 2004).

Secondly, validation for its use of measuring gait in the real-world is limited. This is because measuring real-world gait is a very complex exercise due to the number of contextual factors that may introduce high or low frequency noise to the recorded signals, which limits the applicability of the algorithms introduced in the previous section. Additional complex factors arise from the variety of uncontrolled sources that influence the gait outcome measures, including disease characteristics, patient daily habits, the context in which the gait is recorded and the purpose of the walking. All these factors limit the validity of existing algorithms developed to quantify targeted DMOs.

1.5 Aim of the Thesis

This PhD thesis sits within Mobilise-D, a large European wide project focussed on developing and implementing a digital mobility assessment solution to demonstrate that real-world digital mobility outcomes (DMOs) can successfully predict clinical outcomes and assist clinical decision making. To ensure the digital mobility assessment is suitable in assessing a broad range of mobility levels and adaptations, Mobilise-D aims to validate this digital mobility assessment in five patient populations where measures of mobility loss can present in specific changes linked to the pathology and in all cases can be related to disease progression: Chronic Obstructive Pulmonary Disease (COPD), Multiple Sclerosis (MS), Parkinson's Disease (PD),

Proximal Femur Fracture (PFF) and Congestive Heart Failure (CHF). This will be achieved via two separate stages: a technical validation, aiming at establishing accuracy and reliability of selected DMOs, followed by a clinical validation aiming at proving their construct validity, predictive capacity, and ability to detect change.

Within this context, the specific aim of this thesis is to define, validate and deploy, as part of the above technical validation, an experimental protocol and the relevant algorithmic and laboratory tools needed to enable a robust multicentric collection of gait data from a wide range of disease groups.

The above aim will be achieved through the completion of the three following main objectives:

1. Assess and review previous protocols that aim to test the technical readiness of a wearable sensor as a real-world digital mobility assessment, via a scoping review (Chapter 2).
2. Identify and develop the software and tools needed to ensure reproducibility of experimental procedures in a multicentric study and verify the reliability and robustness of the above tools (Chapter 3).
3. Define and validate a laboratory-based experimental protocol to validate the algorithms used for calculating DMOs. This will be achieved by evaluating the effectiveness of the protocol and looking into its ability to induce a large variety of walking speeds, effectively mimicking real world gait (Chapter 4).

2 CURRENT APPROACHES USED IN ASSESSING THE TECHNICAL READINESS OF A DIGITAL MOBILITY ASSESSMENT FOR REAL-WORLD GAIT: A SCOPING REVIEW

2.1 Introduction

Wearable motion sensors, such as an inertial measurement unit (IMU) or magneto inertial measurement unit (MIMU), are relatively cheap, easy to use and are not restrictive in application or environment, therefore making them a promising solution for assessing real-world gait.

The ability to continuously monitor within a participant's habitual environment may give a fuller picture of the burden a pathology has on a person's day-to-day life and may indicate changes in function that could be related to the onset of disease or disease progression prior to clinical symptoms. However, wearable motion sensors are not currently available as a clinical application to disease monitoring due the restricted understanding of their technical validity as well as the clinical meaningfulness of their outputs when considering a tool that can fully assess and alert to the above factors. A primary reason for this is the complexity that comes with technically validating this form of technology, this can be sensitive and differ for a number of factors such as, device specifications, configuration, modality and algorithms used to quantify the DMOs. In addition, any form of validation would need to be completed on each population of interest that presents with gait patterns that are disease specific. As a result of this complexity, a framework to validate real-world monitoring using IMUs is still lacking.

The aim of this review is to map the current state of protocols that aim at assessing the technical readiness of a sensor solution to estimate real-world walking. Common concepts and possible gaps in the methodology will be identified to shape the framework of a comprehensive protocol, to be defined in the subsequent chapters. To ensure an extensive assessment of the current state of the art, this review will be inclusive of all adult populations and will focus on five main themes that are crucial elements of a technical validation conducted in this context:

- **Wearable sensor used:** The sensor type, configuration of the sensors body location and fixation method.
- **Sample population:** Sample cohort(s) and size included in the data collection.
- **Experimental procedure:** Environment and types of activity used to evaluate the validity of their device/assessment (i.e., experimental procedure).

- **Reference system:** The system used as a reference for corroborating the accuracy and precision of the estimated DMOs.
- **Chosen metrics and Statistics:** The specific gait metrics (DMOs) chosen, the algorithms used to estimate the chosen metrics and the statistical analysis used.

2.2 Methods

This scoping review was conducted in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analysis Statement extension for Scoping Reviews (PRISMA-ScR) (Tricco et al., 2018). As this scoping review would be conducted by only one rater, it was not deemed appropriate to consider the results sound for publication due to possible bias. Therefore, the review was not registered and only used to consider possible themes and gaps in the literature that may guide the authors decision making in the methods developed as part of this thesis.

2.2.1 Article Selection

Technical and clinically relevant databases were searched using keywords and search strings. The initial search of PubMed, Scopus, Web of Science, IEEE and Embase was completed in August 2019. To verify the potential of newly published articles, a second search using the same search strategy was completed, prior to the final definition of the technical validation protocol described in this thesis, in January 2020.

In line with framework of scoping review, the search strategy was purposely designed to be broad across the research area to try and ensure all necessary articles were included in the screening process. The search strategy used focused on screening titles and abstracts and was defined as: *(validation OR reliability OR accuracy) AND (technology OR wearable OR sensor OR inertial measurement OR inertial sensor OR IMU OR IMUs OR MIMU OR MIMUs OR accelerometer OR gyroscope OR insole) AND (walk OR walking OR gait OR mobility) AND (real OR unsupervised OR free OR home OR public).* The appropriate search term notation technique was applied for each database.

2.2.2 Study Selection

Mendeley reference manager was used to remove duplicate citations and screen titles and abstracts based on specific inclusion criteria: (1) it was a full-length journal article originally published in English (2) the article assessed the estimation of walking in human adults (3) there was more than one subject (4) the sensor used was wearable and included at least one

of the components commonly found in an IMU (i.e., accelerometer or gyroscope) (5) at least one spatial or temporal gait parameter was computed using the wearable sensor (6) the computational method was described or referenced (7) there was a comparison of the computed gait parameters against a reference system. Next, the full text assessment was undertaken with the same inclusion criteria implemented. As the analysis of the literature was conducted as a scoping review, no quality assessment, beyond the initial inclusion criteria, was required (Munn et al., 2018).

2.2.3 Data Extraction

Data extraction was then performed on the included articles and focused on the five themes defined above, considering: sample population included, a detailed description of the experimental procedure, type of wearable sensor used, sensor location, the reference system used, the computed gait parameters, computational methods used, and statistical analysis used.

2.3 Results

2.3.1 Search Results

A total of 2,975 articles were identified with the database breakdown as follows: PubMed (450), Scopus (525), Web of Science (1,184), IEEE (357) and Embase (459). When duplicates were removed 2,079 articles remained. After screening of titles and abstracts following the inclusion criteria, 295 articles were included in the full text assessment. Finally, a total of 103 articles were established (Figure 2.1). Due to the large number of articles included, the results and discussion will only reference papers of specific interest, with the full list of articles available in appendix One and the Excel sheet used for the data extraction shared as an additional file (ChapterTwo_DataExtraction.xls).

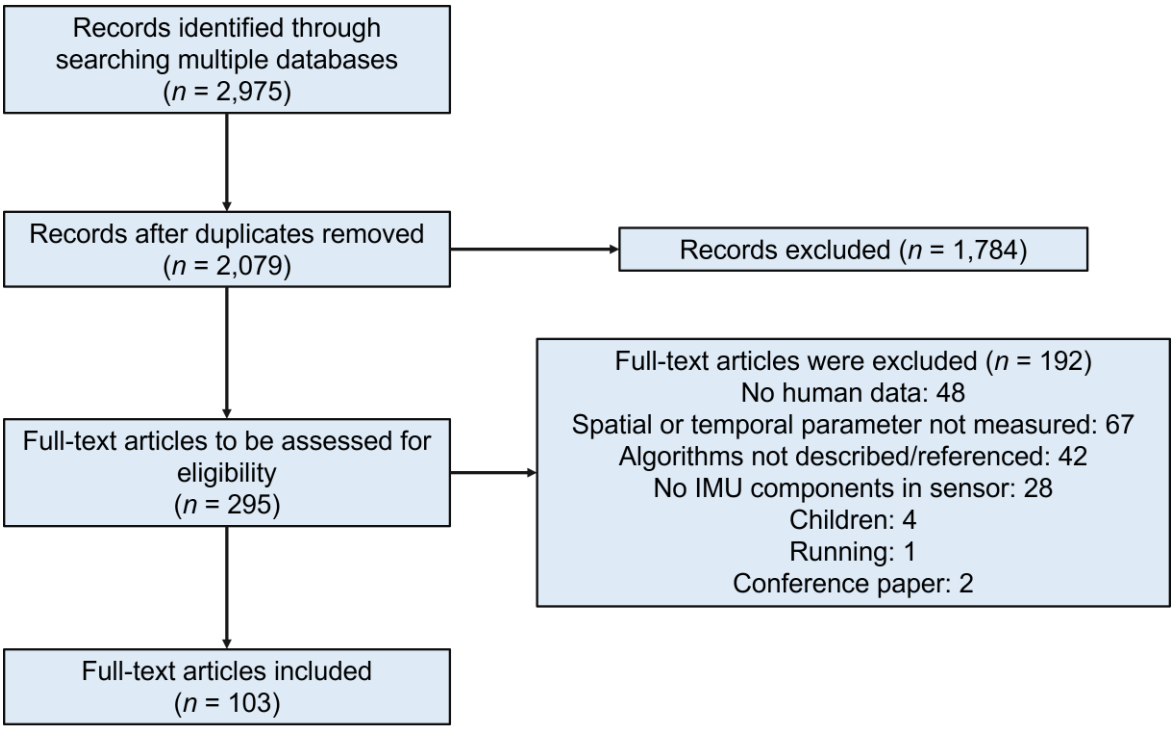


Figure 2.1. Flow diagram of how records were screened for the review.

2.3.2 Wearable Sensor Used

As determined by the inclusion criteria, all wearable sensors included at least an accelerometer or gyroscope, with 45 articles using only accelerometers, 9 using only and 48 articles integrating both accelerometers and gyroscopes in an IMU or in some instances with the addition of a magnetometer in the form of an MIMU. The final article included, used IMUs and the addition of a distance sensor (Bertuletti et al., 2019).

In consideration to the configuration of the sensor, it was evident that there was a large variation across studies, with a total of 10 different locations on the body identified (Figure 2.2) with differences in fixation methods such as; adhesive tape (Godfrey et al., 2015), belts (Schimpl et al., 2011), clips (Ji et al., 2018), a pendant (Brodie et al., 2016) and hearing aids (Atallah et al., 2014). Regardless, a common concept was the use of more than one wearable sensor and location, as found in 53 of the articles included.

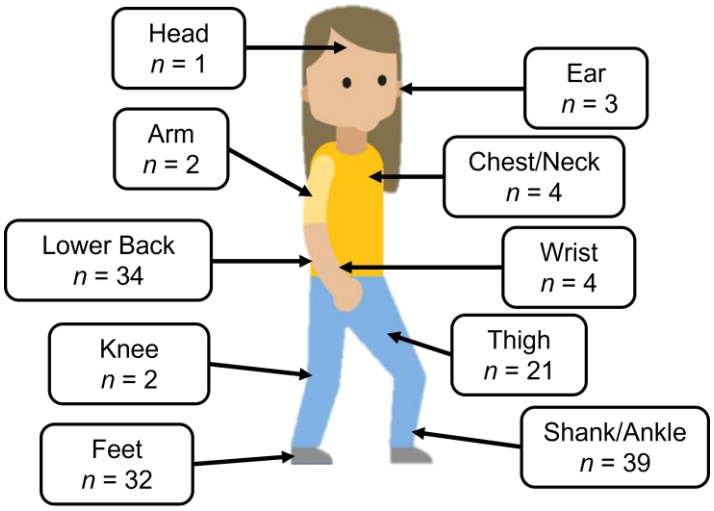


Figure 2.2. Summary of sensor locations found.

2.3.3 Sample Population

Sample size from the 103 articles varied from three participants (Cho et al., 2018; Fourati, 2015) to 116 (Rampp et al., 2015). Sample population also varied across studies, with 71% solely recruiting healthy adults and the other 29% of articles including participants from clinical cohorts. The clinical cohorts included dispersed across six different patient populations and ranged from participants with transtibial amputations (Miyazaki, 1997; Selles et al., 2005), participants who had suffered a stroke and/or hemiparesis (Bejarano et al., 2014; Derungs et al., 2018; Dobkin et al., 2011; Hsing-Cheng et al., 2016; Motoi et al., 2012; Punt et al., 2014; Wang et al., 2018; Yang & Li, 2012), participants with heredity spastic paraplegia (HSP) (Martindale et al., 2020; Prateek et al., 2020), participants with a form of osteoarthritis (OA) (Aminian et al., 1999; De Vroey et al., 2018) and participants with a neurological pathology: either multiple sclerosis (McGinnis et al., 2017; Storm et al., 2018; Supratak et al., 2018) or Parkinson's Disease (Aich et al., 2018; Silvia Del Din, Alan Godfrey, Brook Galna, et al., 2016; Ferrari et al., 2015; Han et al., 2009; Hsing-Cheng et al., 2016; Hundza et al., 2013; Lee et al., 2018; Mariani et al., 2013; Pham et al., 2017; Salarian et al., 2004; Sejdic et al., 2016; Sprager & Juric, 2018; Yang et al., 2011)(Figure 2.3).

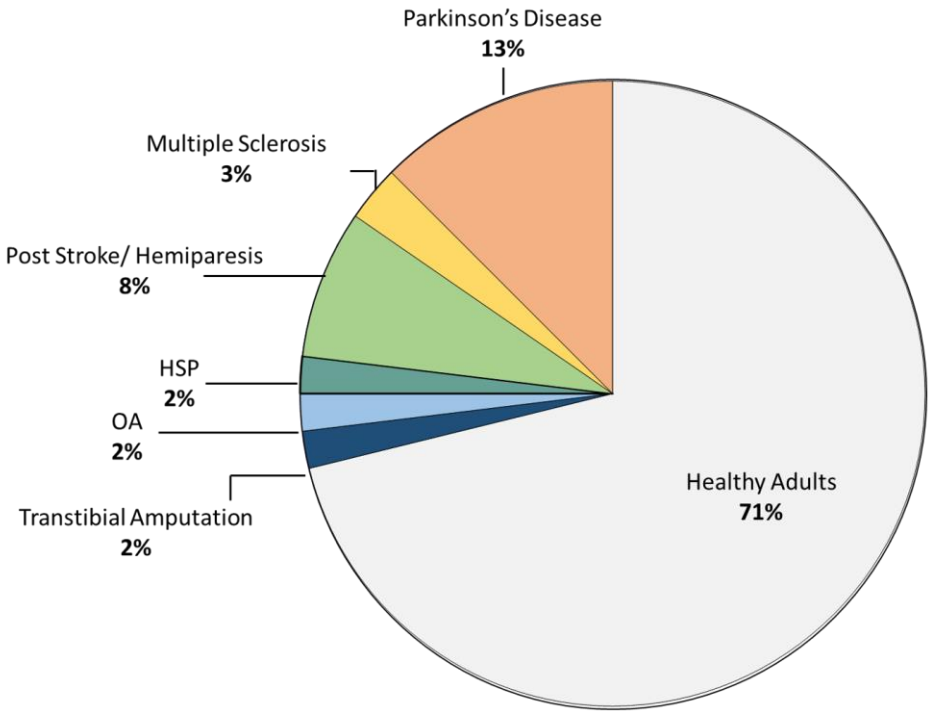


Figure 2.3. Subdivision of the subject populations included in the reviewed articles.

2.3.4 Experimental Procedure

When assessing the environments where the studies were completed, 97% were within a clinic, research setting, or a public space selected by the researcher. The three articles that conducted the data collection in the participants habitual environment, included the nursing home lived in by the participants (Derungs et al., 2018), or following the daily routine of the participant in public spaces and their home (Hickey et al., 2016; Storm et al., 2018).

Four distinct types of task protocols were identified across the included articles (Figure 2.4): a standard straight walking assessment (52% of articles), a set of structured tasks (36% of articles), a semi-standardised trial (10% of articles) and data collection in a free-living environment (2% of articles).

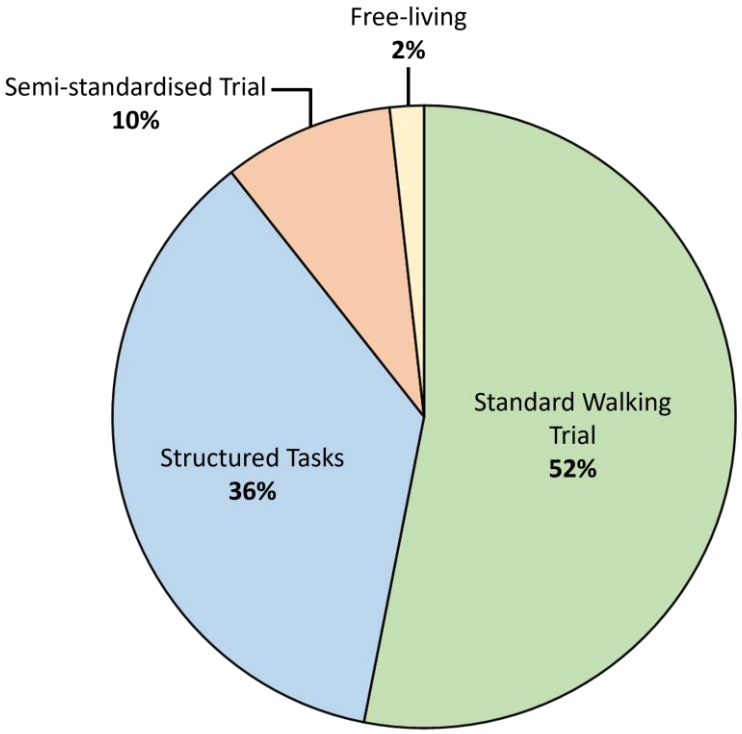


Figure 2.4 Subdivision of the type of task protocols identified.

Straight Walking

Fifty-three of the 103 articles found performed an algorithm validation using only straight walking trials at the participant's self-selected speed. For 43 of these articles, the task was performed overground in a controlled environment with the walkway length varying from 4m (Byun et al., 2016) to 40m (Zhang et al., 2015) across protocols, as well as variation in repetitions of the task, with 27 of the articles only performing one trial and the others performing up to five repetitions (Zhang et al., 2015). For the other 10 articles found, straight walking was imposed via a treadmill, with the speed altered to fit a comfortable walking speed

for the participant. Again, the length of these trials varied across studies from 30 seconds (Ben Mansour et al., 2015) to 6 minutes (Zhang et al., 2019).

Structured Tasks

Thirty-eight of the studies included in this review considered the application of algorithms for the estimate of DMOs in more complex tasks than standard straight walking. Based on the consortium of physical activity (Ainsworth et al., 1993), alteration to walking patterns more realistic to the real-world context could include change in walking speed, incline/steps, surface, path shape and changes in cognitive demand. These were all common alterations in the structured tasks found, however, protocols primarily focused on a single variation, with most focusing on changes in speed to the perceived "self-selected" speed of a standard straight walking trial (Figure 2.5). It is worth noting that all articles, bar three (Mariani et al., 2013; McGinnis et al., 2017; Teufl et al., 2018), also included a standard straight walking test.

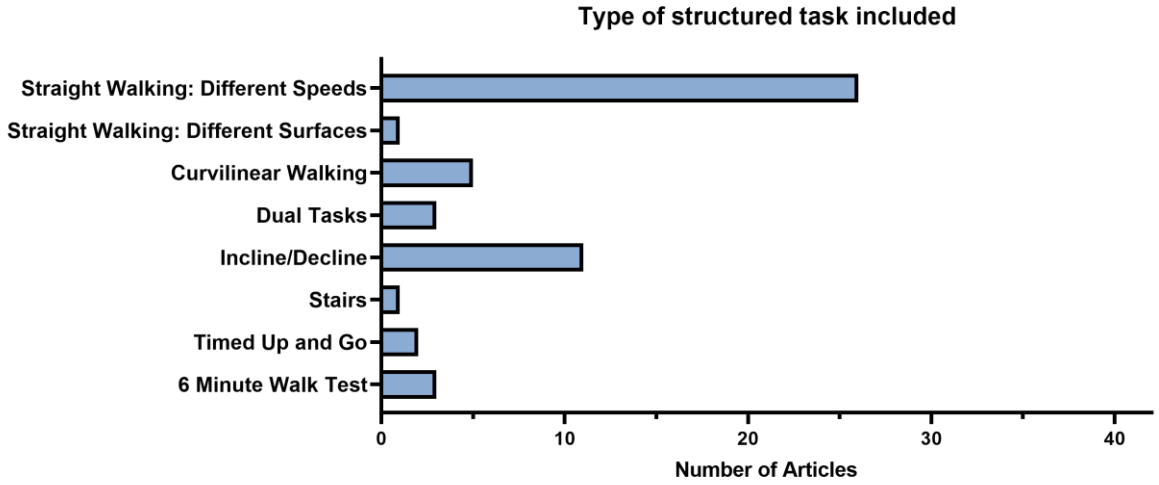


Figure 2.5. Bar graph of the type of structured tasks and frequency found across the included articles.

A Semi-standardised Trial

From the 103 articles included, only 10 incorporated some form of semi-standardised trial in their experimental procedure(Brodie et al., 2016; Derungs et al., 2018; Fasel et al., 2017; Figueiredo et al., 2018; Fourati, 2015; Genovese et al., 2017; Pham et al., 2017; Soltani et al., 2020; Storm et al., 2018; van Oeveren et al., 2018). For the sake of this review, a semi-standardised trial is defined as a battery of tasks in the form of circuit that incorporates postures and movements expected in the real-world. Semi-standardised trials are typically preformed under supervision in a clinic or lab, however, the main difference to the typical structured tasks is that the participant is allowed to move freely while moving between and completing the necessary tasks. Due this form of trial including multiple tasks in one circuit, analysis was completed to determine if the tasks chosen were consistent across articles.

However, the large variation in what activities were included was evident, with straight walking being the only task incorporated in all protocols (Figure 2.6).

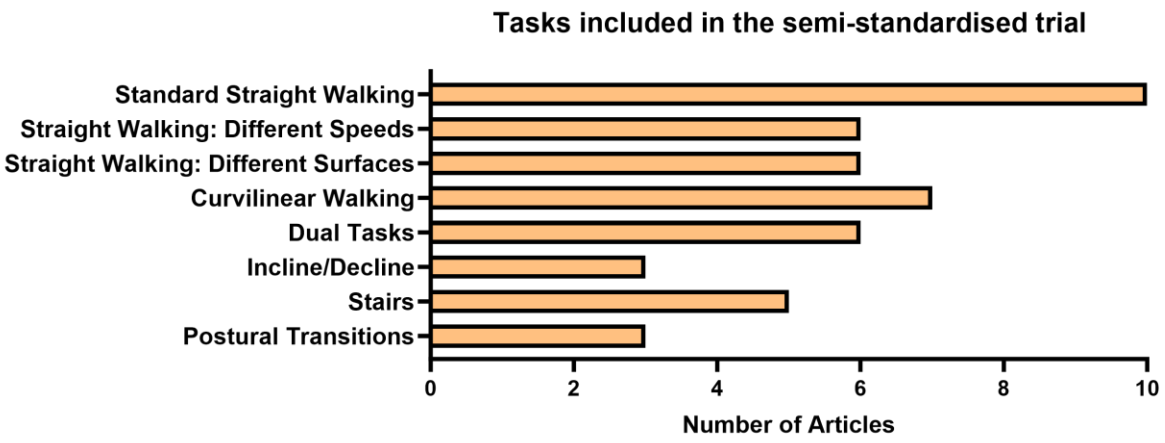


Figure 2.6. Bar graph of the type of structured tasks and frequency found across the included articles.

Free-living

A free-living protocol is defined as a participant performing their daily routine in their preferred environment, without a prescribed protocol or under the supervision of a study investigator. From all the articles included, only two assessed the technical readiness of the wearable sensor solution in this setting (Hickey et al., 2016; Storm et al., 2016). Although both articles included a baseline trial of standard straight walking prior to conducting the free-living data collection, the length of assessment in the free-living environment did vary from 20 minutes (Storm et al., 2016) to one hour (Hickey et al., 2016) .

2.3.5 Reference System

To determine the reliability and consistency of the quantified DMOs, a reference system that provides the level of highest expected accuracy and minimum detectable changes for each selected gait parameter is required. From the articles included, it was evident that there was large variation in the reference system chosen, as show in Figure 2.7.

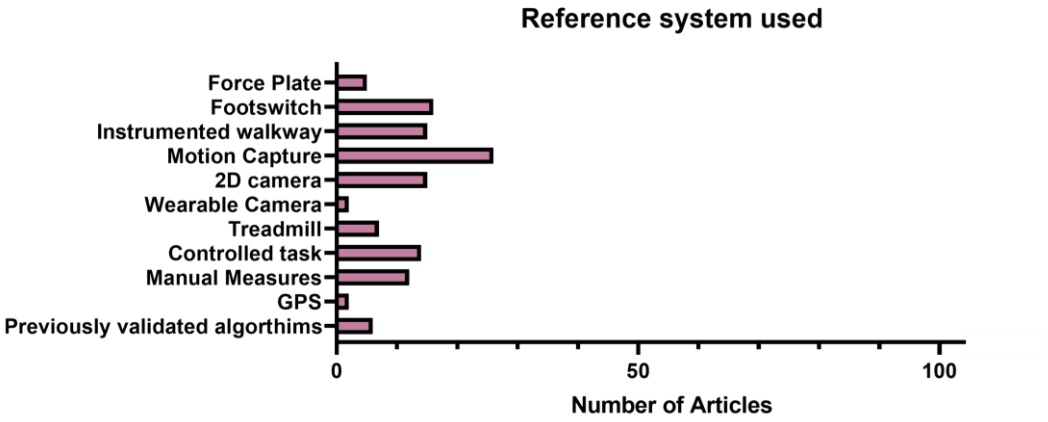


Figure 2.7 Bar chart of the reference systems used in the included articles.

Although all the reference systems identified could be deemed appropriate for the comparison of certain gait metrics, the more sophisticated reference systems, specific to the assessment of human movement, were more commonly used in the included articles:

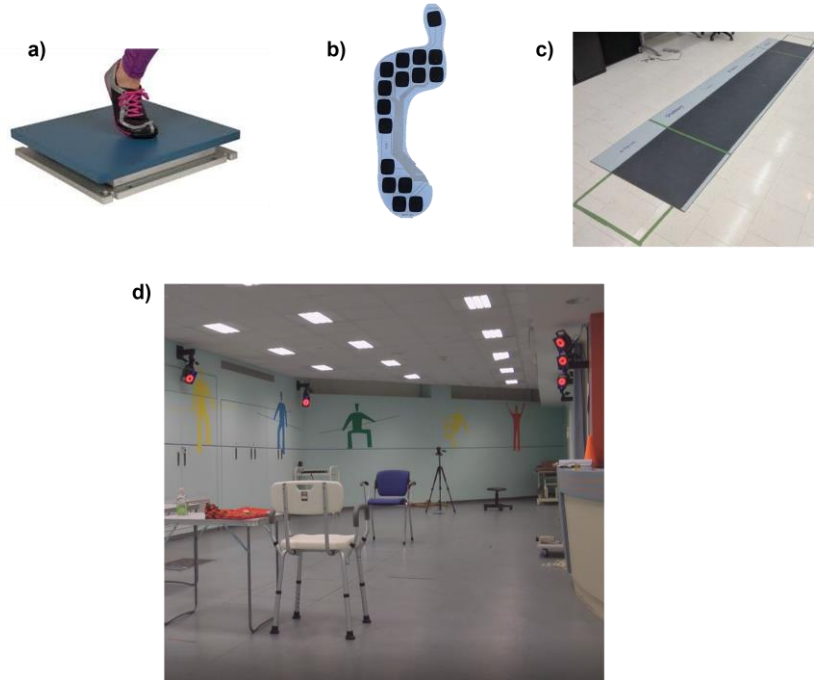


Figure 2.8 Examples of the more sophistacted reference systems used in the included articles: a) force platform b) footswitch or pressure insole c) instrumented walkway d) motion capture system.

Force Platform

Force platforms are instrumented plates usually integrated to be flush with the ground that measure the ground reaction force (GRF) generated during a movement. This equipment is highly sensitive to load, meaning that gait events can be detected with a high level of accuracy making them a common system to use when validating temporal measures of gait.

Footswitch or Pressure Insole

A footswitch or pressure insole can directly detect foot contact with the ground during gait. The most common implementation of a footswitch or pressure insole use force sensitive resistors (FSRs) which are very thin (~1mm) sensors that act as variable resistors and take advantage of the force vs resistance relationship to generate a voltage that is proportional to the exerted force (Lowe & ÓLaighin, 2014). In the instance of a footswitch, the number of FSRs is typically minimal, with one on the big toe and heel of each foot and taped directly to the plantar aspect of the foot or to the insole of the shoe. Pressure insoles typically comprise of more FSRs that map the whole plantar aspect of the foot in the form of an instrumented insole (Djuric-Jovicic et al., 2014),

Instrumented Walkway

Instrumented walkways consist of a pressure mat that integrate multiple force sensitive resistors, using a similar method as the footswitch. The use of infrared distance sensors is also sometimes incorporated. A distance sensor detects the presence of nearby objects by emitting a beam of infrared and detecting a change in the return signal. In relation to their use in an instrumented walkway, the infrared sensors run horizontally across the mat, to accurately detect the foot at the initiation of gait events, in particular ICs.

Optoelectronic Stereophotogrammetric System

Optoelectronic stereophotogrammetric (SP) systems estimate the 3D coordinates of an object based on two or more 2D photographic images taken simultaneously by cameras with different views of the object. To ensure a clear image of the object, SP systems use a combination of charge-coupled device (CCD) cameras to record light and retroreflective markers that are attached to the object to identify the specific points of interest or different segments of the object (e.g., joints and body segments in human movement analysis) (Robertson et al., 2014)

2.3.6 Chosen Metrics and Statistics

Chosen Metrics

From the 103 articles included, 55 solely focused the validation of temporal based metrics such as gait events, stride time and cadence. A further 13 articles focused on validating spatial based metrics such as stride length, step length and step width, with another 18 focussing the validation on the spatiotemporal based metric, walking speed. The remaining 17 articles considered metrics from all three domains.

From an algorithm perspective, it was evident that although variations were made in all articles in regard to sample population, context, change in pre-processing of the sensor components (such as filtering or sensor fusion) or investigation of bespoke or adaptive thresholds, the fundamental approaches for calculating the gait parameters were similar. For the temporal parameters of gait this was primarily using the method of feature extraction (71 of the included articles), double integration for the spatial parameters (34 of the included articles) and integration of these two outputs for the spatiotemporal measures (21 of included articles).

Algorithms for Estimation of Temporal Parameters

Two main algorithmic theories for the estimation of the temporal parameters were found, with both using features of the sensors signals to identify the events and phases of gait for different locations on the body: the lower back and the lower leg:

***Lower Back*:** This approach derives from a method proposed by Zijlstra and Hof (2003) that predicted acceleration signal characteristics based on the modelled vertical and forward displacement of the centre of mass (COM) during the gait cycle based on their previous study (Zijlstra & Hof, 1997). The model was able to identify that during single support there is an increase in forward acceleration of the COM due to the propulsion of the contralateral leg going through initial swing, followed by a resultant downward fall in the vertical acceleration as the contralateral leg passes during midstance. Subsequently, as the gait transitions from single to double support, the forward movement of the COM decelerates and the vertical acceleration changes into an upward movement. The proposed algorithm identified the peak forward acceleration prior to the zero-crossing of the signal (transition from single to double support) as an IC. From validation of the algorithm against ground reaction force (GRF) data the results showed small errors in event time (Figure 2.9).

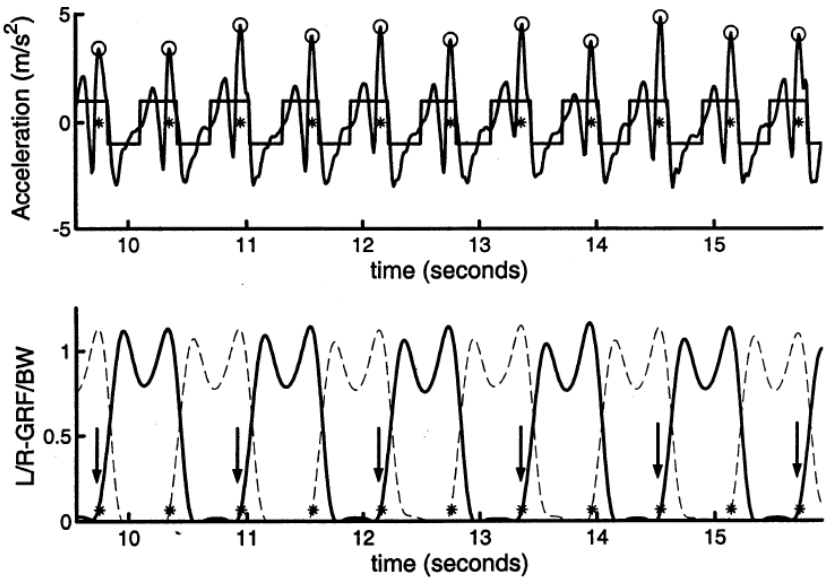


Figure 2.9. The upper figure shows the forward acceleration signal from the lower back as a solid black line. The lower figure shows right (solid black line) and left (dashed line) GRFs. In both figure the asterisk indicate the IC as detected from the left and right GRFs. The open circles indicate the FC detected from the acceleration data using the proposed method (Zijlstra & Hof, 2003).

This algorithm has since been refined (Zijlstra, 2004), however, this approach has its limitations with irregularities in the acceleration signal sometimes resulting in the methods inability or incorrect detection of events (González et al., 2010).

An alternative approach to identifying gait events using vertical acceleration is based on processing the signal through continuous wavelet transforms (CWT). By removing superfluous signal fluctuations, the underlying frequency variations can be observed, allowing appropriate detection of time related events in the signal. This method was developed by McCamley et al.(2012) and applied a Gaussian CWT on the vertical acceleration signal, with the local minima then corresponding to IC. After further CWT differentiation, FCs were identified as the jerk maxima (Figure 2.10).

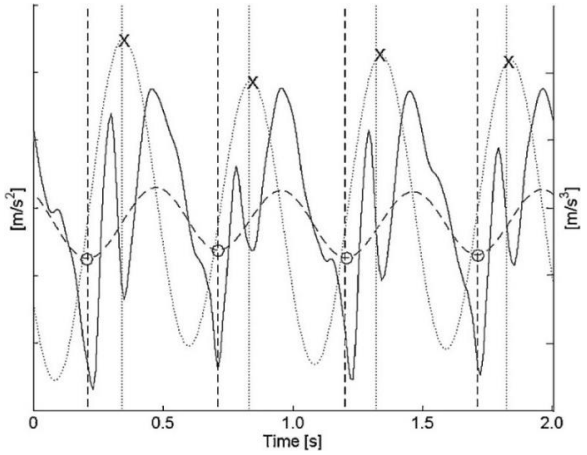


Figure 2.10. The vertical acceleration (solid line) is integrated and then differentiated using CWT (dashed line). The minima from this signal corresponds to the ICs (o). Further CWT differentiation (dotted line) provides the jerk maxima which corresponds to the FCs (x) (McCamley et al., 2012)

On validation, this method was able to detect 100% of events when compared to GRF data. Evidently this method was able to demonstrate its robustness and consequently has been implemented within both articles that performed a validation in a free-living environment, of which a variety of healthy and patient populations were included (Hickey et al., 2016; Storm et al., 2016).

***Lower Leg*:** This concept is based on the clear identification of mid-swing, due to flexion of the knee at this stage of the gait cycle resulting in a consequential high peak in forward angular velocity. The local minima prior to this peak is considered FC and the local minima after is considered IC (Salarian et al., 2004) (Figure 2.11). This method has been widely used in various subject demographics in semi-structured trials and free-living environments (Martinez-Hernandez & Dehghani-Sanij, 2018; Qiu et al., 2016; Storm et al., 2016; Zhou et al., 2016).

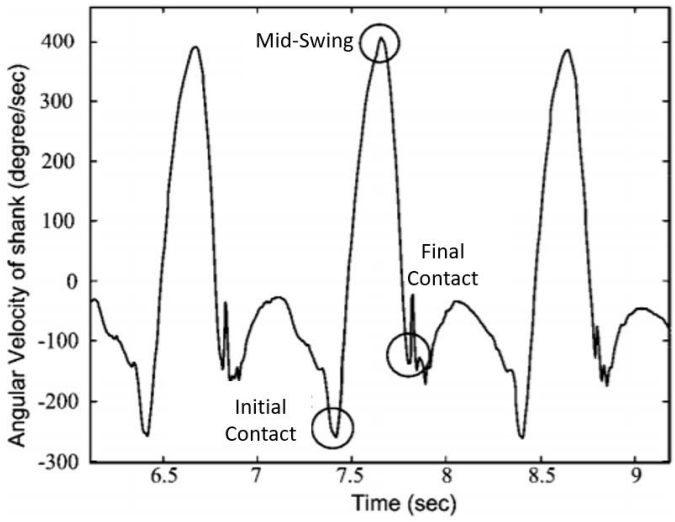


Figure 2.11. Feature extraction from gyroscope positioned on the shank (adapted from (Salarian et al., 2004))

A foot mounted gyroscope can detect gait phase using the characteristics of the forward angular velocity signal. Theoretically this signal should be zero during foot flat, however a more appropriate approach is to define when the signal becomes approximately constant. Using the detection of foot flat the IC can then be considered as the first instance in which the angular velocity is within a range empirically determined close to the null angular velocity and FC considered as the global minima (Figueiredo et al., 2018) (Figure 2.12).

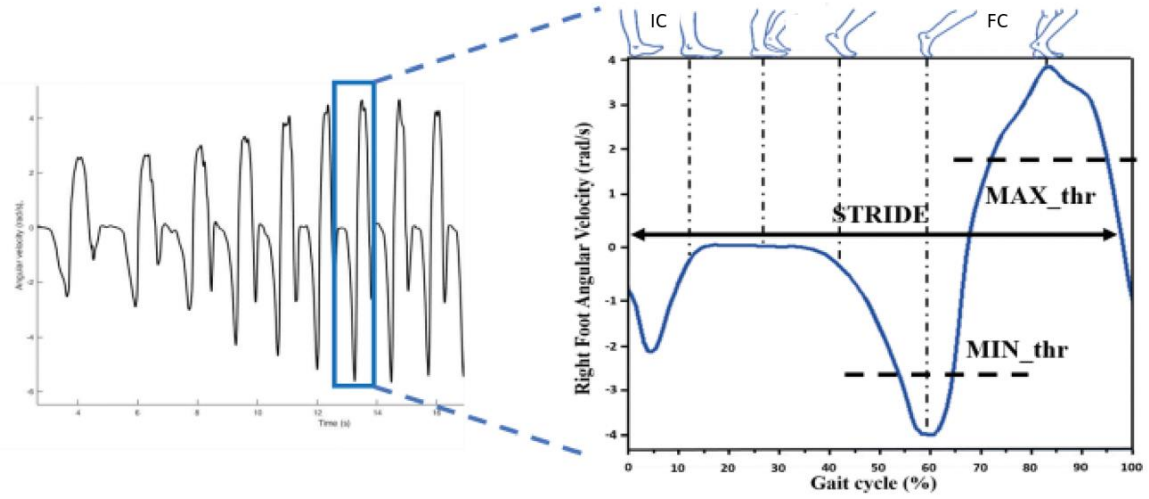


Figure 2.12. Feature extraction from gyroscope positioned of the foot (adapted from (Figueiredo et al., 2018))

Algorithms for Estimation of Spatial Parameters

From the included articles, 23 estimated spatial parameters of gait using numerical integration of accelerometer data from the feet or shank. Using this method, the change in displacement of the sensor between specific gait events of the same foot/leg (conventionally the ICs) can be calculated as an estimate of stride length. To avoid propagation of errors from inherent noise included in the inertial sensor and time varying offset which can lead to an accumulating drift, articles that used this approach included a zero-velocity update (ZUPT), which offsets any drift or noise when the foot is static during mid-stance (Skog et al., 2010).

A further 15 of the included articles estimated stride length by quantifying to displacement of accelerometery data on the lower back. As first described by Zijlstra and Hof (2003), this model assumes that the centre of mass rotates as an inverted pendulum about the ankle during the single support at stance phase, followed by horizontal displacement during the double support at stance phase. Therefore, the stride length can be estimated from the vertical displacement of the COM and the leg length.

Algorithms for Estimation of the Spatiotemporal Parameter

The primary method found for estimating walking speed utilises the ratio between the estimated step or stride length and time, previously described in the sections above. The calculation of this ratio can be defined in multiple ways, including walking speed at step, stride or even walking bout level. The flexibility in calculating this measure can offer opportunity to

explore the spatiotemporal feature of gait within domains such as variability, rhythm and pace (Costa et al., 2022). However, to allow comparison of results, it is necessary the approach used is well defined within the protocol. In the instance of articles included in this review, only six of the 17 articles that estimated the spatiotemporal metric defined the equation used for walking speed (Ferrari et al., 2015; Hannink et al., 2017; Hao et al., 2019; Hundza et al., 2013; O'Brien et al., 2019; Salarian et al., 2004).

Statistics

As the validations focused on comparing the DMOs estimated from the sensor to the reference system, all articles chose a set of statistics that could highlight the differences in this comparison. This included, the mean absolute error, relative error, limits of agreement, root mean square error and correlation. Although all articles presented good results that were deemed appropriate for future application of their method, the variation in statistics lead to heterogeneity in the results and limits confident comparability between studies.

2.4 Discussion

With the aim of mapping common concepts and possible gaps in the approaches used for assessing the technical readiness of digital mobility assessments for real-world gait, this review considered five specific themes crucial to the development of a protocol in this topic. From the data of the 103 articles included, it was evident that there were a range of common concepts within these themes, however, their execution led to a general level of heterogeneity across the literature.

2.4.1 Wearable Sensor Used

Although the inclusion criteria restricted the review to measuring systems that contained traditional sensor components of an IMU, it was evident that there was a split between IMUs and sensors with individual components. Although the constraint to a single sensor component may restrict the implementation to certain algorithms, this did not seem to hinder the efficacy of the measure and subsequent validation. As such, this exemplified the versatility these wearable sensors have, with minimal restrictions to their application regardless of cost differences, integration in commercially available devices or as an experimental device made for the research purpose.

The large variation in sensor configuration and fixation does limit comparability of results between the literature, however, it again emphasises the versatility of these devices. In

application of any device (medical or commercial) the acceptability and wearability are crucial. With the range of sensor locations and modes of application available, there is potential to tailor measures to the user's preference and comfort as well as the desired outcome, only strengthening the possibility for clinical application. However, the variations in signal features at different locations of the body and the possible inclusion of noise from certain fixation methods (Küderle et al., 2022; Niswander & Kontson, 2021) does currently limit the comparability of data, even in the locations more frequently used such as the lower back.

2.4.2 Sample Population

With the high number of articles included using a healthy adult population it is a likely suggestion that this is due to easier access to this subject group, both in regard to ethics and general availability of participants within a research environment. In addition, the overall physical health of this population allows the opportunity to assess complex movements and extended durations of continuous walking while not over burdening the participant or compromising their safety. Furthermore, this limits optimisation of the algorithms in dealing with any deviations to the recorded signals that may occur in the presence of disease specific gait. However, this limits the findings of these articles, with many highlighting the need for further validation in patient cohorts prior to application.

Of those articles that did include one of the six patient cohorts identified, limitation was found in the sample sizes and repeated studies within the cohort, with only the cohort of Parkinson's Disease being included in over 10 of the included articles. As previously mentioned, validation for each clinical cohort is crucial in future application of monitoring real-world gait due to disease specific gait features and subsequent changes in the recorded sensor signals (Silvia Del Din, Alan Godfrey, Brook Galna, et al., 2016). Although it is understandable that smaller sample sizes and less studies including clinical cohorts is likely, it is crucial that results from these studies is comparable to aid meta-analysis and strengthen the position of their application within these populations.

2.4.3 Experimental Procedure

Although this review included articles that aimed to complete a validation for real-world application, it was evident that 89% of studies did not include an aspect of real-world within their task protocol, with only 40% of these studies including an additional task that was not straight walking. A primary reason for this is likely the inclusion of some deviation in the study from the other literature (i.e., differences in sample population or algorithm). As such, a

baseline measure of structured walking was the most effective in identifying errors in the methods used without incurring external errors from contextual factors or fluctuation in walking patterns seen within different real-world scenarios. Although the need for a baseline measure proved to be necessary, with all articles that performed some form of real-world gait also including a straight walking task, it is evident that this assessment is inadequate to validate use in real-world due to the aforementioned factors. Consequently, further validation in more complex tasks and environments would be needed.

From those articles included that performed a structured task, out with standard straight walking, and those that completed some form of semi-standardised trial, the activities included varied. The lack of agreement between the tasks set in each study could be due to the validation only being completed on specific characteristics of gait such as cadence (Hwang et al., 2018) or characterising walking during stair negotiation(Zhou et al., 2016). Although the restriction to specific tasks is a fair approach in validating the system for its primary purpose, it does limit the results of the validation to the specific tasks included in the protocol, which when considering a validation for a method assessing real-world gait is not sufficient.

With only two articles including a validation fully immersed in the participants real-world (i.e., performed in their habitual environment with no restriction) it is evident that this form of assessment is the most complex in the context of a technical validation. This is likely due to the constraints in reference systems available for real-world, as discussed below, as well as considerations needed in determining what is the most appropriate setting for assessing real-world gait (e.g., home, work, public space) and how long must the assessment be to capture all necessary activities to provide meaningful data for a statistically sound validation, while not overburdening the participants.

2.4.4 Reference System

Reference systems again varied between articles, with use of standard measures (e.g., known distances) and more sophisticated systems specific to the assessment of human movement. When determining which reference system to use it necessary to determine, which DMOs will be assessed and the capability of the systems to output reference outcomes, the compatibility of the system to the environment, cost and availability of the system, and the expertise needed to process reference outcomes from the system.

Standard measures of known distance, time, and manual dictation of steps/gait events either by eye or 2D video can offer a simple measure to compare but does not come without its limitations. As previously mentioned, this form of measure is limited to the basic parameters of gait and does not allow assessment of the more discrete parameters of gait that are likely to fluctuation in the real-world. In addition, as these measures rely on manual assessment,

reliability can vary, and output of the reference outcomes can become burdensome on the assessor in the case of manual dictation of steps. Regardless, a benefit to this from of reference system is the minimal restriction to environment, aiding its popular use in the semi-standardised trial and free-living protocols (Brodie et al., 2016; Derungs et al., 2018; Hickey et al., 2016; Pham et al., 2017).

The more sophisticated measures used, all referred to as gold standard solutions (i.e., the best available measures) in the assessment of human movement, can produce high accuracy for a variety of gait measures but do not come without limitations. This is due to most gold standards solutions consisting of technologies that confine the experiments to a laboratory location or to constrained movements (i.e., straight walking with instrumented walkways and treadmills or confinement to lab environment for motion capture systems and force platforms). Although 46% of articles included used one of these gold standard solutions, their use was confined to the structured task and standard straight walking protocols.

Although there are a select few gold standard solutions that are wearable and can be used with no constraints to environment, including footswitches or pressure insoles (Aminian et al., 2002; Aminian et al., 1999; Figueiredo et al., 2018; Han et al., 2019a; Hsing-Cheng et al., 2016; Jarchi et al., 2014; Lee et al., 2010; Lee & Park, 2011; Maqbool et al., 2015; Martinez-Hernandez & Dehghani-Sanij, 2018; Misu et al., 2017; Sabatini et al., 2005; Storm et al., 2016; Zhou et al., 2016) and global positioning systems (GPS) (Barnett & Cerin, 2006; Soltani et al., 2020), these systems again can come with the limitation only measuring a select set of gait parameters, as described for the standard measures. To combat these current restrictions, seven of the articles included, used additional IMUs or MIMUs (usually on the shank or feet) that had previously been validated for the purpose of gait. However, due to the possible propagation of errors incurring from use of a reference that is not considered a gold standard solution, this method requires an extensive validation prior to use and has currently been limited to studies within their own research group (Bertuletti et al., 2019; Silvia Del Din, Alan Godfrey, & Lynn Rochester, 2016; Fasel et al., 2017; Storm et al., 2016; Storm et al., 2018; van Oeveren et al., 2018; Vincenzo et al., 2017). This limits its appeal as a reference system however the variety of DMOs this solution could offer is beyond that considered with the current gold standard solutions. To this end, one article considered the combination of both a wearable gold standard and IMUs (Bertuletti et al., 2019), allowing a greater understanding of all the calculated DMOs while still considering the integrity and high accuracy of the reference system.

Regardless of technology used, the integrity of the reference systems measure needs to be considered and defined within the validation protocol. This results in a necessary

standardisation of the system across participants and processing steps, both of which were missing information within the described protocols.

2.4.5 Chosen Metrics and Statistics

Based on the above restrictions for reference systems and task protocols, it is of no surprise that over half of the articles included solely focused on validation of temporal gait metrics, with all but one article (where a GPS was used as reference for walking speed (Soltani et al., 2020)) validating this measure in semi-standardised protocols. The inclusion of spatial metrics was always conducted in a laboratory environment, making full use of the gold standard solutions such as the motion capture systems and instrumented walkways. However, this resulted in all but three articles (Mariani et al., 2010; O'Brien et al., 2019; Wang et al., 2015) only assessing forms straight walking, again limiting the applicability of the solutions implementation in the real-world. This assessment focus to straight walking was also seen in all articles that included assessment of walking speed.

What did become apparent, was the common use of certain algorithmic approaches for estimation of the desired DMOs. This general agreement in the fundamental methods used in extracting gait parameters from wearable sensors, has aided in the optimisation of the original algorithms, as shown by the strongly positive results described in the articles included. However, limitation to the comparability of results is present due to the variation of statistics used and general heterogeneity in the other areas of the validation protocol.

2.5 Conclusion

The review included 103 articles that aimed to validate wearable sensors for the estimation of real-world gait. Based on the abundance of literature found and the versatility of the measuring tool regarding sensor components, body location and fixation, the potential of wearable sensors as an appropriate tool for clinical application is clear. However, the general lack of agreement in protocol used limits comparability of the results. Based on the five themes, recommendations from current concepts and limitations from the literature can be made to framework future protocols that assess the technical readiness of a wearable sensor solution to monitor real-world gait:

- Due to the abundance of literature already available for healthy adult populations, sample populations should focus on a variation of clinical cohorts (Silvia Del Din, Alan Godfrey, Brook Galna, et al., 2016), with healthy adults used as a control to any new method with sample size being determined based on statistical power.

- Sensor components used and their metrological characterisation should be well defined within the study to allow translation of methods that do not restrict to cost or specific commercial devices. With location and fixation of sensor also defined and considered when comparing to results from previous literature.
- Experimental protocol should focus on a tiered approach, from standard straight walking to more complex structured tasks and finally assessment in a real-world context. Structured tasks should make use of the available gold standard solutions for reference and incorporate a large variation of complex gait that focuses on scenarios found in the real-world (Ainsworth et al., 1993).
- Duration and context of a real-world validation should be considered and defined to determine the meaningfulness of the data and validate its acceptability regarding participant burden.
- Gold standard solutions should be the focus in laboratory-based validations with standardisation of their use clearly defined and transparency of the systems or experimental errors described and minimised where appropriate.
- Use of a multi-sensor system that includes wearable gold standard solutions is recommend for real-world validation, however, extensive validation is required prior to its application as a reference system.
- Gait metrics chosen should aim to produce a holistic validation of gait parameters (i.e., temporal, spatial and spatiotemporal) to incorporate all aspects of real-world gait. Algorithm validation should focus on the strong agreement and results already produced in the literature, however, a clearly defined and standardised statistical should be presented and experimental data should be shared to allow comparison to future studies.

3 ENABLING THE USE OF AN OPTOELECTRONIC STEREOPHOTOGRAMMETRIC SYSTEM AS A GOLD STANDARD IN A MULTICENTRIC STUDY

3.1 Introduction

The term “gold standard” is a long-standing definition in the world of medicine and science, with first reference in a 1975 Lancet Review (Jones & Podolsky, 2015). A gold standard in terms of clinical or technical validation and development is the use of the best available measure to determine the accuracy and capability of a new intervention or tool. Although the term “gold standard” is widely used within research, it is crucial that the gold standard status of a measure be challenged through rigorous testing, to ensure that this is truly the best measure available and attain a comprehensive understanding of its limitations (Versi, 1992).

Optoelectronic stereophotogrammetric (SP) systems are integral in the field of human movement research for quantifying kinematic variables, through the instantaneous 3D tracking of retroreflective or light emitting markers via high performance optical cameras (Baker et al., 2018). With their use spanning three decades, the continuous evaluation of the accuracy and robustness of these systems, and optimisation with the advancement of technology have resulted in their recognition as the gold standard for the validation of technologies with a similar purpose.

Regardless, the gold standard status should be wavered with caution due to the multiple routes for error and risks associated with the use of a SP system. Errors can be characterised as either random error (i.e., the level of error that comes from the capability of the system, such as the level of noise in the 3D reconstruction of the markers) or as a systematic error (i.e., the level of error accumulated from the system setup, environment, and experimental design). Although random errors from SP systems have been widely characterised in the literature (Table 3.1), a standardised approach to characterise or mitigate sources of systematic error is yet to be fully established. Subsequently, the level of data quality across operators and laboratories can vary, which in turn may limit the comparability of data and possibly compromise the gold standard status of the measure.

Most systematic errors occur during the calibration (i.e., before the data collection), in the attachment and model calibration of the marker set used (i.e., during the data collection) or during the pre-processing that is required prior to calculation of outputs (i.e., after data collection). This chapter will describe the errors in detail for each timepoint mentioned above and propose and validate the methods used in the multicentric study to minimise error and ensure comparability between systems.

Table 3.1: Performance of SP systems in terms of accuracy as reported in the literature from 1997 – 2017 (adapted from (Conconi et al., 2021))

Paper	OSS	Measure adopted	Accuracy definition	Accuracy value	Maximum error definition	Maximum error value
(Ehara et al., 1997)	Vicon 140, 4 cameras	Known distance	Mean absolute error	1.60 ± 1.82mm	Maximum distance variation	10.87mm
	Vicon 370, 6 cameras			0.94 ± 0.39mm		12.94mm
	Ariel APAS, 2 cameras			11.61 ± 5.36mm		37.54mm
	Dynas 30/h, 2 cameras			18.42 ± 0.24mm		78.68mm
	ELITE PLUS, 4 cameras			0.53 ± 0.31mm		2.15mm
	Expert Vision, 4 cameras			1.14 ± 0.53mm		12.22mm
	PEAK5, 2 cameras			3.85 ± 2.04mm		18.49mm
	PRIMAS, 2 cameras			1.79 ± 0.14mm		10.23mm
	Quick MAG, 2 cameras			2.25 ± 0.52mm		14.58mm
	Video Locus Color, 2 cameras			7.63 ± 2.81mm		41.81mm
	Video Locus Reflective, 2 cameras			7.73 ± 1.45mm		35.42mm
(Richards, 1999)	Vicon 370	Known distance	Root mean squared error	0.62mm	Max absolute error with no more than 3 cameras	5.57mm
	Ariel APAS			4.27mm		4.94mm
	Chamwood CODA			4.87mm		9.26mm
	BTS Elite Plus			1.73mm		16.13mm
	Motion Analysis HiRes			0.59mm		5.99mm
	Qualisys ProReflex			0.80mm		12.76mm
	Peak Perform. Motus			0.91mm		5.82mm
(Miller et al., 2002)	Motion Analysis, six cameras	Known displacement	Mean absolute error	0.05mm	–	–
(Windolf et al., 2008)	Vicon-460, 4 cameras	Known displacement	Root mean squared error	63 ± 5µm	Maximum grid-point error	416 ± 129µm
(Kuxhaus et al., 2009)	Vicon M-612 ViconPeak, 6 cameras, Vicon Workstation v4.5	Known displacement	Mean absolute error	0.05 ± 0.005mm.	Maximum absolute error	3.7mm
(Yang et al., 2012)	Vicon MX, F40 cameras, 5 cameras, Nexus 1.6.1	Known displacement	Mean absolute error	<2µm	Mean absolute error	<7µm
(Diaz Novo et al., 2014)	Vicon MCam-60, 8 cameras, Vicon-Workstation V4.6	Known distance	–	–	Maximum mean error	<5mm
	Vicon T160, 12 cameras, Vicon Nexus 1.7					<5mm
	Canon Zr300, 3 cameras, Hu-m-an V5					<20mm
(Eichelberger et al., 2016)	Vicon Bonita, 6 or 8 or 10 cameras, Vicon Nexus 1.8.5	Known distance	Mean error best case	0.08 ± 0.05mm	Mean error worst case	2.30 ± 0.001mm
(Aurand et al., 2017)	OptiTrack Prime 41, 42 cameras, OptiTrack Motive 1.10.1 Final software	Known displacement	Root mean squared error	<200µm on 97% of the volume	Root mean squared error, worst case	<1mm
(Di Marco et al., 2017)	Vicon system MX-series, 8 cameras, Vicon Nexus 1.8.5	Known distance	3*STD of mean absolute error	0.1mm	Maximum root mean squared error in dynamics	0.4mm
	Vicon system T-series, 10 cameras, Vicon Nexus 1.8.5			0.3mm		1.7mm
(Merriaux et al., 2017)	Vicon T40S, 8 cameras	Static- Known displacement Dynamics- Known distance	Static-mean absolute positioning error	0.15 ± 0.015mm	Dynamics-Maximum error	<2mm

3.2 Before the Data Collection: Establishing Random and Systematic Errors

A substantial part of the material presented in this section has been published in:

Scott K, Bonci T, Alcock L, Buckley E, Hansen C, Gazit E, Schwickert L, Cereatti A, Mazzà C, on behalf of the Mobilise-D Consortium. A Quality Control Check to Ensure Comparability of Stereophotogrammetric Data between Sessions and Systems. Sensors. 2021; 21(24):8223. doi: 10.3390/s21248223

Please see Appendix Two for author declarations.

3.2.1 Background

SP systems estimate the 3D coordinates of an object based on two or more 2D photographic images taken simultaneously by charge-coupled device (CCD) cameras with different viewpoints of the object. As CCD cameras are light sensitive, SP systems utilize this feature to distinguish between ambient light and retroreflective or light emitting markers that are attached to the object. To amplify the markers brightness in comparison to skin, clothing, and background, most CCD cameras have light emitting diodes mounted around the lens (Robertson et al., 2014)(Figure 3.1).

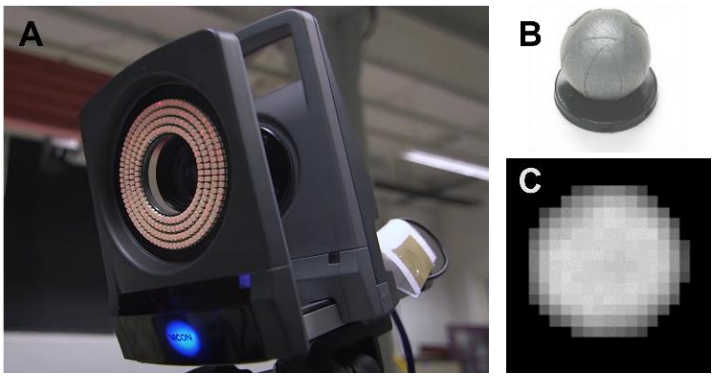


Figure 3.1Overview of SP equipment; a) a CCD camera from a SP system with the light emitting diodes mounted around the lens, forming the red ring b) photo of a SP retroreflective marker c) Example of the 2D image of the marker captured by the CCD camera.

To construct the 3D spatial coordinates of the markers from the 2D images, the 3D coordinate reference system of the camera is projected in the 2D image planes. Using a method of direct linear transformation (Abdel-Aziz & Karara, 2015), the 3D position of the marker is calculated while solving a system of equations that consider the intrinsic and extrinsic parameters of the camera. The inclusion of these parameters are crucial, with the intrinsic parameters allowing for any corrections due to distortion to the captured image from possible fault in the lens (i.e., misalignment between the lens and optical axes of the camera), and the extrinsic parameters

defining the position of the camera to a global reference coordinate system (GCS), which consequently allows images of the same marker from different cameras to be combined, constructing the 3D position of the marker within the 3D space/laboratory (Figure 3.2).

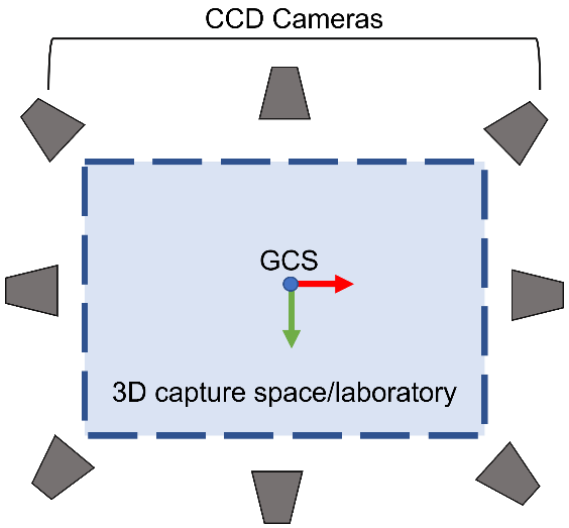


Figure 3.2 Schematic of a typical SP system layout, with cameras surrounding the desired 3D capture space.

To ensure a high level of accuracy in the reconstruction of the markers 3D position, both the intrinsic and extrinsic parameters need to be estimated prior to any data collection via a system calibration, which consists of two parts. Both procedures require the use of a rigid object that contains a set of markers, where the geometrical relationship of the markers are already defined within the system. In the initial calibration, the object is moved through the 3D capture area to exploit the total volume seen by the cameras. During this part of calibration, the system will define the intrinsic parameters based on deformations of the calibration object in the images captured and will determine the cameras relative position to the others. Secondly, the calibration object is placed level at the location determined as the GCS (ideally on the ground in the centre of the capture space). Based on the position and rotation of the object, the GCS is then set.

Manufacturers of SP systems commonly report that if these system calibration procedures are performed correctly, SP systems can track 3D marker trajectories with submillimetre accuracy and precision (van der Kruk & Reijne, 2018). However, in practice, factors such as the number of cameras, camera resolution, camera positioning, laboratory environment and capture volume can affect the accuracy of these systems (Chiari et al., 2005). As a result, rigorous testing to attain a metrological characterisation and comprehensive understanding of the consistency in error between sessions and across systems is crucial for protocol designs such as, repeated measures, longitudinal studies and multicentric studies.

Although SP systems have an internal quality assessment of the random and systematic errors during the calibration procedures, the reporting varies between manufacturers and does not entail the provision of easily readable feedback. In response, standard procedures to metrologically characterise different SP systems, by quantifying the error in marker reconstruction, have been proposed. However, these methods can be complex and regularly require the use of bespoke equipment, with a cluster of markers attached to turning plates (DeLuzio et al., 1993; Merriaux et al., 2017; Richards, 1999), sliding plates (Aurand et al., 2017), rigid rods (Della Croce & Cappozzo, 2000; Diaz Novo et al., 2014), sliding blocks allowing adjustable linear movement of the marker cluster (Everaert et al., 1999; Miller et al., 2002; Windolf et al., 2008), or articulated arms with 3 degrees of freedom (DeLuzio et al., 1993) and regularly involve a time intensive assessment in respect to the number of recordings required (3 – 45 per check) and modifications in equipment setup between each trial (Aurand et al., 2017; DeLuzio et al., 1993; Diaz Novo et al., 2014; Everaert et al., 1999; Miller et al., 2002; Richards, 1999; Windolf et al., 2008). Although the quantification of random errors is regularly reported in these methods, few have considered the quantification of systematic errors. Often this has been limited to a dynamic capture in the centre of the capture volume consisting of rotational movements (Della Croce & Cappozzo, 2000; DeLuzio et al., 1993; Diaz Novo et al., 2014; Merriaux et al., 2017; Richards, 1999), which may not represent the systematic errors accumulated in dynamic movements typically seen in human movement data.

In response, recent studies have considered the addition of dynamic checks more representative of human movement data. To this end, Eichelberger et al. (2016) used a plate consisting of two markers attached to the foot, knee, and sacrum during a straight walking trial to determine the error at the three most common heights of marker placement throughout the chosen volume of capture. Although this study clearly demonstrated the need for assessing systematic errors under dynamic conditions, the proposed setup requires a separate trial for each position of the plate and subsequently would involve additional time prior to each data collection, which may not always be available in clinical-based laboratories. Additionally, this approach did not consider errors in the estimates of angles between multiple markers, relevant when estimating angular kinematics. More recently, Di Marco et al. (Di Marco et al., 2017) proposed a quick dynamic check of systematic errors using the calibration object provided by the SP manufacturer as it moved through the capture volume for approximately 20 seconds. Their method led to similar results in a participant-based check (a maximum error of 0.7° for the object compared to 2.4° for the gait trials) however, validation of this method is restricted, since the reliability of the check across different systems, operators, and calibration objects has not yet been assessed. Additionally, automation of the processing and reporting of errors was not implemented and its ability to

ensure comparability as a routine quality control (QC) check in SP data collection across different laboratories has not been assessed.

Based on the encouraging results from Di Marco et al. (2017) this study aims to develop and verify the reliability and usability of a simple and time effective QC check to estimate the random and systematic errors of different SP systems. To ensure the QC check can be adopted in the routine running of SP data collections, three main criteria will be taken into consideration during its development; there will be minimal burden to the operator, no need for additional equipment and automated reporting of random and systematic errors. To determine the QC check's reliability, robustness and ability to discriminate a change in systematic error between calibration/session, a verification procedure will be completed to determine whether the QC check is a) reliable, regardless of the calibration object used and b) robust regardless of operator and system calibration/session. Additionally, to demonstrate its ability to ensure the comparability of SP data between systems, the proposed QC check will be deployed in a multicentric study that consists of five laboratories with varying SP systems and operators.

3.2.2 Methods

The Quality Control (QC) Check

The proposed QC check makes use of the manufacturer's calibration object (or a similar alternative, as described later in the methods) during two short acquisitions. The calibration object chosen for this study was the Vicon calibration wand (Vicon Motion Systems, Oxford – UK) (Figure 3.3), which will be referred to as *CO#1* throughout. *CO#1* was used for all data collection except the assessment of the QC check's reliability with a different calibration object.

To first isolate random errors associated to the internal SP systems algorithms in marker reconstruction, the calibration object is positioned in the middle of the capture volume and a static trial of approximately 5 seconds is recorded. With this placement, the "ideal" capture can be quantified with minimum influence from calibration outcome, camera placement or intrinsic camera settings (Aurand et al., 2017). Subsequently, to quantify the systematic errors accumulated by the specific system setup and calibration outcome, the calibration object is moved through the capture volume at a velocity comparable with that used in the dynamic phase of the system calibration, for a minimum of 20 seconds or until the full capture volume has been covered.

Verification of QC Check Reliability

The reliability of the proposed QC check was assessed through a verification procedure that included assessment of the three variables subject to change on its use: the calibration object used (Part a), the operator performing the trials and the system calibration/session (Part b).

Data for assessing characteristics of the calibration object was collected using a 10 camera Vicon system T-Series (camera model: T160, camera resolution: 16mpxs, capture volume: 6.0mx4.0mx2.1m) with a sampling frequency of 100Hz and processed in Nexus 2.8.2 (Vicon Motion Systems, Oxford – UK). Prior to data collection, the system was calibrated with a minimum of 3000 frames successfully capturing the calibration object for each camera. All testing was completed at a single site, on the same day and same calibration.

As SP manufacturers calibration objects can vary between the use of retroreflective (passive) and light emitting (active) markers, the effects of using either marker type was assessed. To ensure consistency in marker configuration and object shape, two versions of *CO#1* were used to complete the QC check. A single run of the QC checks was completed with the active marker version of *CO#1* (Figure 1A) and then repeated using the passive marker version of *CO#1* (Figure 1B).

Prior to quantification of the errors, the calibration objects marker configuration and known inter-marker distances and angles must be defined. For *CO#1*, these geometrical relationships include both linear and angular measures (Di Marco et al., 2017), considering both the shortest and longest marker distances (Figure 1). The known inter-marker geometries for this object were based on the measures given by the SP manufacturer.

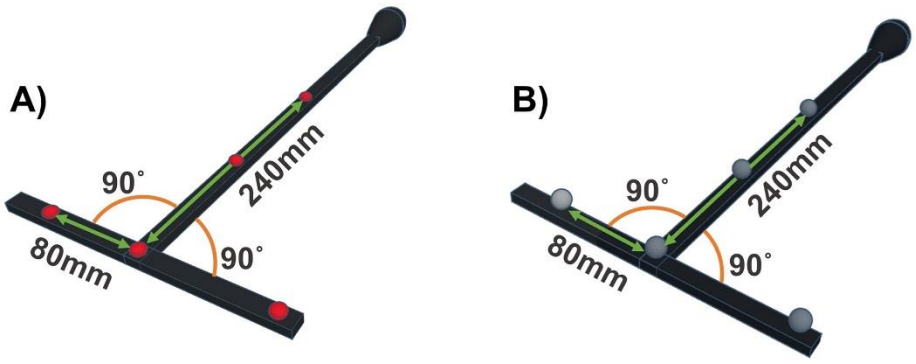


Figure 3.3. Schematic of the defined linear and angular marker geometries of CO#1. A) the active marker version of CO#1 B) the passive marker version of CO#1

To determine the robustness of the QC check when using different calibration objects and marker configurations (Part b), the QC check was also performed using *CO#2*. *CO#2* was a modified version of the passive marker version of *CO#1*, which consisted of two additional passive markers fitted to replicate the inter-marker distances in different locations of the object (Figure 2A). The additional markers were attached manually with 3 repeated measures using a calliper to ensure accurate positioning. Furthermore, a 3D printed object that contained three passive markers was rigidly attached to the top of the object using stronghold tape, to determine the adaptability of the QC check to calibration objects with a third dimension in the marker configuration (Figure 2B). The dimension of the 3D object was specified in the manufacturing and as a sound check, the dimensions were then manually assessed using a repeated measures of a calliper and goniometer. Using this approach should ensure the same level of accuracy as the dimensions of the original object from the manufacturer.

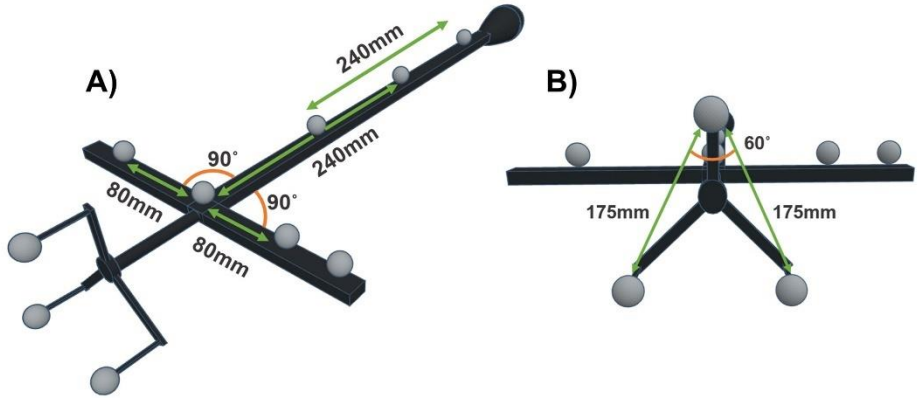


Figure 3.4. Schematic of the defined linear and angular marker geometries of CO#2 A) defined marker geometries of the 2D marker configuration B) defined marker geometries of the 3D marker configuration.

To ensure the reliability of the QC check over repeated measures, determine if the QC check can identify a change in systematic errors between system calibrations/sessions and demonstrate the robustness of the QC check when performed by operators with varying levels of expertise, three conditions were considered: intra-operator intra-session, intra-operator inter-session, and inter-operator intra-session, respectively. All data were collected using the same SP system, specifications, and calibration procedure described above. The two intrasession conditions were completed with different system calibrations and on separate days. The active version of *CO#1* was used for all data collections (Figure 3.3A).

The intra-operator intra-session condition was completed by the same operator (OP): *OP#1*-highly experienced with the QC check procedures. Three repetitions of the QC check were

performed during the same system calibration/session. The intra-operator inter-session was again completed by *OP#1* with one recording completed on three different system calibrations/sessions. Finally, the inter-operator intra-session condition was completed by three different operators during the same system calibration/session. To evaluate the ease of performing the QC check and reliability of the outcome, the three operators had varying knowledge and experience with the check: *OP#1, OP#2*– had a good understanding of the QC check but did not perform it regularly, *OP#3* – had no prior knowledge or use of the QC check.

All operators were instructed to perform the trials using the same description and language as stated:

Static Trial: "Please place the calibration object level on the floor in the middle of the capture volume and record a trial of the object in this position for 5 seconds."

Dynamic Trial: "Please move the calibration object at a velocity comparable with the one you would use in the system calibration procedure for at least 20 seconds. Please make sure to exploit the full volume of the desired capture area."

Multicentric Deployment

To evaluate its suitability in the context of a multicentric study, the QC check was implemented as part of the data collection for the IMI project Mobilise-D (Mazzà et al., 2021). This study includes data collected from five different SP systems in different locations, with varying laboratory and system setups, SP system manufacturers and operators, as shown in Figure 3.5. The desired capture volume to be covered by all SP systems was defined as 5mx4mx2m. Each site was instructed to calibrate its SP system following their standard procedures. Prior to implementation, all operators were trained by *OP#1* on how to perform the QC check using the instructions described in the inter-operator intra-session protocol. All sites used the *CO#1* object for this study (Figure 3.3), with either active or passive markers. Each laboratory completed the QC check on 10 different system calibrations all on different days of data collection.

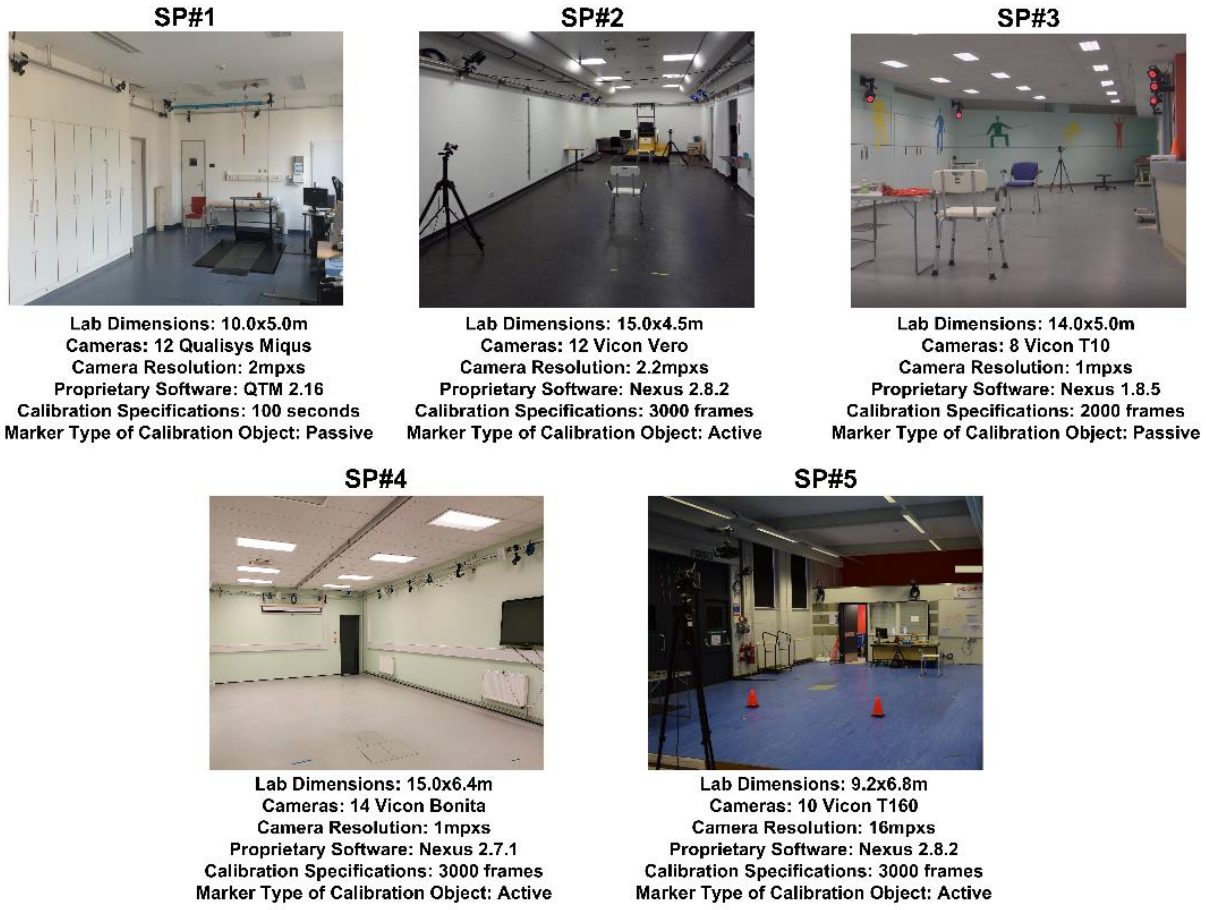


Figure 3.5. Specifications of the five SP systems used in the multicentric deployment.

Data Processing and Analysis

All data was reconstructed and labelled using the manufacturer software and recommendations. As adaptability to varying SP manufacturers was desired, a c3d file format was chosen for data export due to its universal use by different SP software. Using the calibration object's marker trajectories, the distances and angles of the reconstructed markers were quantified. The error between the reconstructed and known inter-marker geometries defined above (Figure 3.3 and Figure 3.4) were calculated and the random and systematic errors were quantified as follows:

To characterise random errors accumulated from the SP marker reconstruction, the standard deviation of the error from the static trial was used to quantify the expanded uncertainty. By selecting a coverage of $k = 3$, coverage of 99.7% of the random errors for a given session was obtained:

$$Expanded\ Uncertainty = SD_E \times k.$$

The systematic error of the dynamic trial was calculated as the root mean square error (RMSE) of the difference between the known inter-marker geometries (y), as defined in Figure 1 and Figure 2, and the corresponding reconstructed inter-marker geometries ($\hat{y}$) for each frame of capture (i) over the full trial (N):

$$\boldsymbol{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (\boldsymbol{y_i} - \boldsymbol{\hat{y}_i})^2}{\boldsymbol{N}}}$$

To allow for immediate reporting of the QC check results, a graphical user interface (GUI) was designed in MATLAB 2020a (MathWorks, Natick – USA) that reads the exported c3d files, compiles the calculation of the errors described above into an automatic pipeline and generates a report for straightforward interpretation of the errors by the operator (Figure 3.6). To allow amendment to any object and marker configuration, the base code used in the GUI for the QC check analysis and example data for both trials of the QC check has been made available via FigShare for use by the scientific community and published as supplementary material in the corresponding publication (Scott et al., 2021).

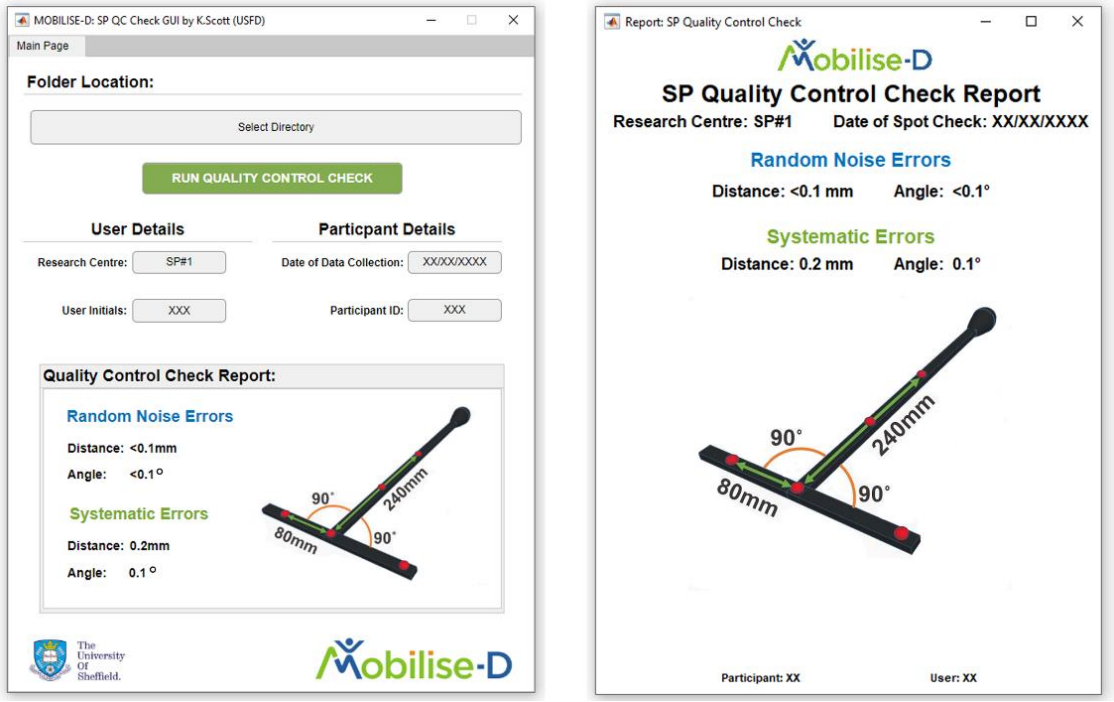


Figure 3.6. QC Check GUI created for multicentric deployment (left) and generated PDF reported (right).

3.2.3 Results

Based on the initial analysis of the random and systematic errors, the quantified error did not show bias to differences in the distances or amplitude of the angle. Therefore, only the highest error for the inter-marker distances and angles are reported in the results.

Verification of QC Check Reliability

In all testing completed for the variation in the calibration object (Part a) the random errors, as calculated by the expanded uncertainty, was below 0.1 mm for the inter-marker distances and below 0.1° for the angles. The systematic errors (RMSE) are reported in Table 3.2, with the highest RMSE for the marker distances of 0.8 mm and 0.5° for the angles.

Table 3.2. The systematic error (RMSE) calculated for the single trial for each of the different calibration objects.

	RMSE			
	CO#1		CO#2	
Measure	Active Markers	Passive Markers	2D Configuration	3D Configuration
Distance (mm)	0.2	0.5	0.6	0.8
Angle (deg)	0.1	0.3	0.5	0.4

In all testing for the variation in operator and session (Part b), the random errors (expanded uncertainty) showed the same results as seen in Part a, with the error of the inter-marker distances and angles always below 0.1 mm and 0.1°, respectively. The systematic errors, as quantified by the RMSE for the dynamic movement of the calibration object are reported in Table 3.3.

Table 3.3. Mean and standard deviation of the systematic error (RMSE) for the three trials for each of the different operator and session conditions.

	RMSE		
Measure	Intra-operator Intrasession	Intra-operator Intersession	Inter-operator Intrasession
Distance (mm)	0.2±0.1	0.3±0.1	1.0±0.1
Angle (deg)	0.1±0.1	0.1±0.1	0.4±0.1

Multicentric Deployment

The random errors (as calculated by the expanded uncertainty) for all 50 QC checks performed across the five SP systems, was always below 0.3 mm and 0.3° for the inter-marker distance and angles, respectively. The systematic errors (RMSE) quantified with the dynamic check are presented in Figure 3.7, with the highest RMSE for the inter-marker distances and angles below 2.5 mm and 2°, respectively.

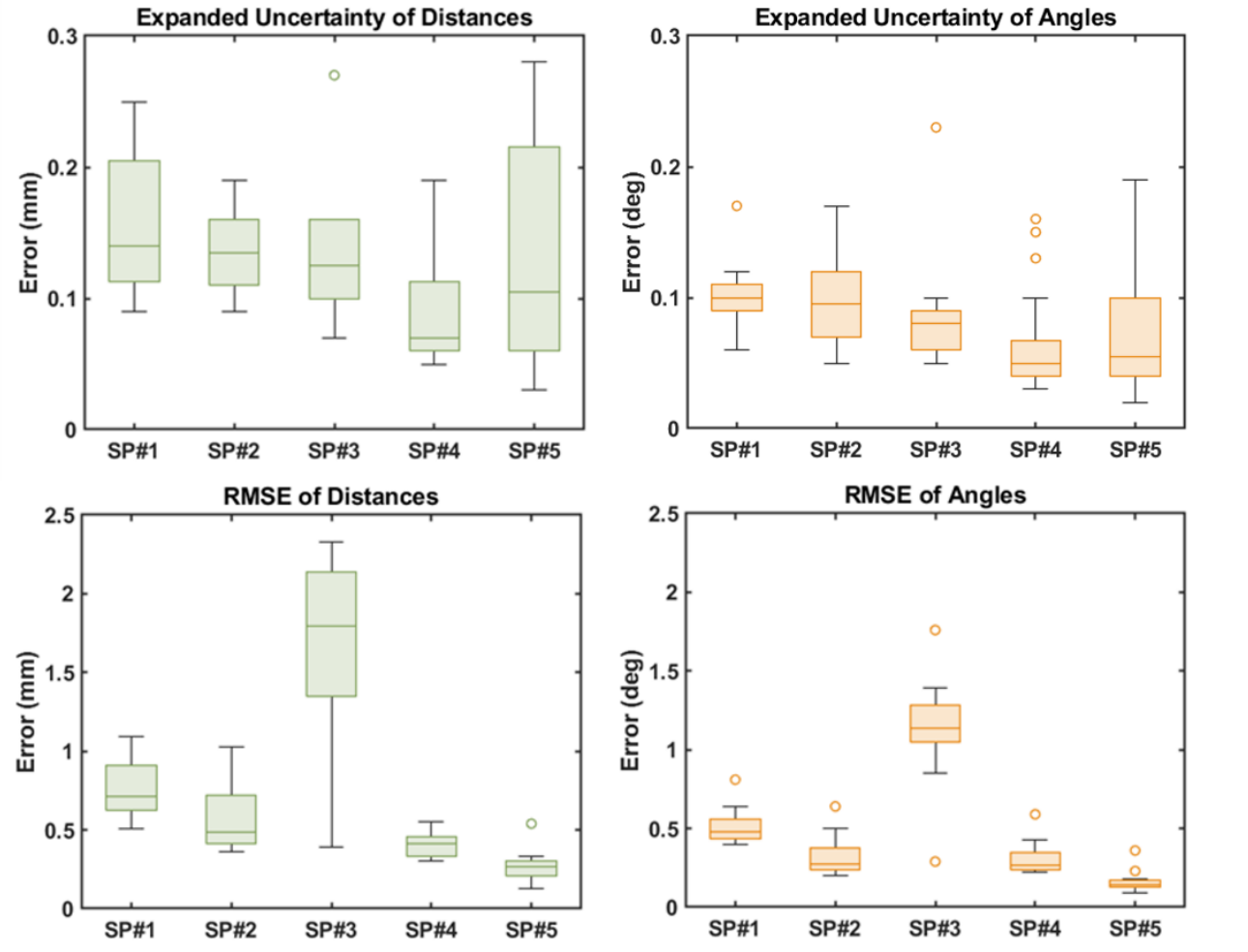


Figure 3.7. Box charts of the expanded uncertainty and RMSE calculated for the distances and angles of the 10 QC checks completed by the five sites

.

3.2.4 Discussion

This study aimed to develop a simple and time effective QC check to estimate the random and systematic errors of different SP systems as part of the routine running of SP data collections. The reported results showed that by using the SP systems calibration object, with

an assessment period of 25 seconds and automated calculation and generation of the errors, the proposed QC check can be successfully completed and interpreted well within 5 minutes. The validated QC check can be performed prior to starting a SP data collection and could be adopted with minimal delays or burden to the operator. The ability to perform such a check with no additional equipment is beneficial due to its wide implementation and routine use within standardised operating procedures to ensure accurate and reliable data collection.

Verification of QC Check Reliability

The QC check was reliable in quantifying random and systematic errors between sessions regardless of the calibration object used, the operator performing the check and the system calibration/session.

The uncertainty quantification showed virtually no changes in the SP systems reconstruction of static inter-marker distances and angles regardless of object, marker type, operator, or session. This demonstrates that using the middle of the capture volume as an "ideal" location for determining the random noise errors associated with reconstruction capabilities of the SP system is suitable and is minimally impacted by the calibration outcome. In addition, the errors quantified are comparable to previous studies that used bespoke equipment in determining random errors in a more structured manner (DeLuzio et al., 1993; Diaz Novo et al., 2014; Everaert et al., 1999; Merriaux et al., 2017; Miller et al., 2002; Richards, 1999; Windolf et al., 2008).

The quantified systematic errors showed minimal change throughout the capture volume covered during the dynamic trials, with all errors at submillimetre and sub-degree level, supporting the gold standard status of the system (van der Kruk & Reijne, 2018). The range of systematic errors reported agree with previous methods that have used more structured assessments to consider the systematic errors at different static points of the capture volume (Aurand et al., 2017) and are comparable with the errors quantified in the previously proposed dynamic checks reported by Di Marco et al. (2017) (1.7 mm for inter-marker distances and 0.7˚ for angles) and Eichelberger et al (2016) (<1 mm error in inter-marker distances).

The slight increase in error observed for the calibration objects with passive markers could be explained by decreased precision of the reconstruction when compared to active markers (Maletsky et al., 2007). Nonetheless, the error quantified fit the defined capabilities of an SP system and therefore can be considered negligible. The systematic errors quantified for the three operators with varying experience showed negligible differences (0.1 mm for distances and 0.1° for angles). Additionally, the QC check accurately identified the systematic errors

related to changes in the system calibration while remaining precise across repeated measures, as shown in the results for the intra and intersessions.

Multicentric Deployment

The deployment of the QC check as part of the multicentric study which includes a variety of systems proved to be successful, with all sites smoothly completing the checks and interpreting the results prior to data collection in a time efficient manner (i.e., within the five-minute window stated above). In addition, the use of the accompanying GUI provided an automatic pdf export of the QC check report that could be appended with the SP data collected, to ensure easy reference and transparency of the systematic errors across sites and sessions.

As shown in the quantification of the random errors for the 50 static trials acquired, the expanded uncertainty for the inter-marker distances and angles were all below 0.3 mm and 0.3° respectively (Figure 5), corroborating the findings from Di Marco et al., (2017). Although the level of random errors observed in the multicentric deployment was found to be slightly higher than in the verification of QCs reliability (maximum difference of <0.2 mm for the marker distances and <0.2° for the angles), the error showed negligible change both within and between SP systems. Moreover, as the calibration object used (*CO#1*) was the same in shape and marker configuration across all five sites (Figure 3.3), the slightly higher quantification of errors is likely due to the variation in camera specifications (e.g., number of cameras, camera resolution and camera placement) as well as the internal algorithms for marker reconstruction varying between the SP manufacturers. This supports the concept of the QC check being able to quantify and isolate random errors specific to a variety of systems.

The precision of the systematic errors quantified across the 10 sessions of data collection for each site showed minimal levels of change, with all reported errors comparable to the dynamic errors reported in previous studies (DiMarco et al., 2017, Eichelberger et al., 2016), with the exception of *SP#3* (Figure 3.7). As the random errors quantified for *SP#3* in the middle of the capture volume fell into a similar range as the other SP systems, a possible reason for the higher systematic error and variation of this error between sessions, is the limitation of having only eight cameras covering the defined capture volume of 5mx4mx2m when compared to the other systems that ranged from 10 – 14 cameras, as well as a smaller camera resolution of 1mpxs. Operation of an SP system with a smaller number of cameras has previously been shown to increase the level of systematic error and decrease the precision of marker tracking (Aurand et al., 2017). Moreover, as the lab size was comparable to two of the other systems used in this study (*SP#2* and *SP#4*), *SP#3*'s higher systematic errors are considered to be primarily due to the limited ability of the fewer cameras to cover the full

capture volume. Nonetheless, for the broader aim of the multicentric study (Mazzà et al., 2021), as spatiotemporal gait parameters are the primary focus from the gait data collected, a maximum linear error of 2.4 mm and angular error of 1.8° is certainly in an acceptable range when scaled to the outputs quantified (e.g., stride length, walking speed and turning angle).

The main limitation of the proposed QC check is the dependency of accuracy in the calibration object manufacturing and the assumption that there has been no deformation of object during standard use. Any inaccuracy of this sort, however, would also affect the system calibration and performance. In addition, due to the SP systems available in this study, only two SP manufacturers have been tested. However, with the source code of the GUI made available and the use of the universal c3d file, it is hoped that other systems could also use this check to produce a wider understanding of errors across different manufacturers.

3.2.5 Conclusion

This study clearly proved that the proposed QC check is feasible to perform in a short timeframe with minimal burden to the operator and has a clear potential to be used as a routine procedure in multisession and multicentric studies. Its adoption in the multicentric study allowed for transparency of error source throughout the validation process and as such has enriched the analysis of our experimental tools.

3.3 During the Data Collection: Standardising a Marker Set and Model calibration

3.3.1 Background

Marker Set

As described in the section above, SP systems require the attachment of retroreflective or light emitting markers to track an object. In the application of this method within human movement analysis, typical biomechanical models simplify skeletal structures of the body as rigid segments that interact with other segments (e.g., the interaction between the thigh and shank). Based on this approach, a 3D local coordinate system (LCS) fixed to each segment can be constructed through the placement of multiple markers and move correspondingly with the participant. As each segment has its own LCS, the multiple LCSs can be related to the GCS of the 3D capture area by means of transformation (Cappozzo et al., 2005)(Figure 3.8) subsequently permitting the quantification of the spatial position of each segment relative to the rest of the body and the kinematics between segments (i.e., the knee joint).

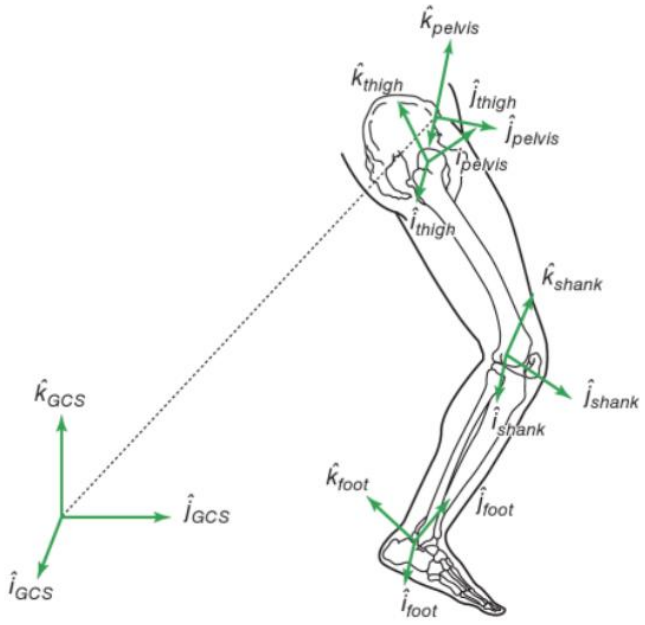


Figure 3.8: The GCS and the LCSs of the right leg (adapted from (Robertson et al., 2014)

The fundamental rule when establishing the LCS of a segment is the minimum requirement of using three markers to define the position and rotation of the 3D orthogonal axes (XYZ). The placement of these markers typically relies on the palpation of anatomical features (e.g., bony landmarks of the skeletal structure such as the epicondyles of the knee) or a virtual marker that is estimated and anatomically scaled based on the location of the anatomically

placed markers (e.g., estimation of the hip joint centre using methods (Davis et al., 1991). While this method is essential in defining the LCSs, it does come with its limitations in assessing human movement due to the morphology of the skin, and consequently the position of the anatomically placed markers in relation to the skeletal structure during activity. To correct for this, tracking markers attached to the relatively flat (and lateral) surfaces of the structure are commonly used to improve the tracking of the segment and LCS throughout a movement.

Although approaches for defining and tracking the LCS for each segment of the body, referred to from now on as marker sets, are readily available within the literature (Cappello et al., 1997; Davis et al., 1991) (Figure 3.9), these marker sets do not come without some considerations prior to use. Most importantly the output of the assessment should be considered, and the marker set refined accordingly to reduce the complexity of the data collection for the participant (i.e., maximise comfort and reduce time of assessment), the operator (i.e., reduce time of assessment and pre-processing of data) and mitigate errors associated with the use of markers (i.e., incorrect placement, loss/noise in the marker trajectory due to cross-talk between markers in close proximity or noise to the marker trajectory associated with soft tissue artifacts (Stagni et al., 2005).

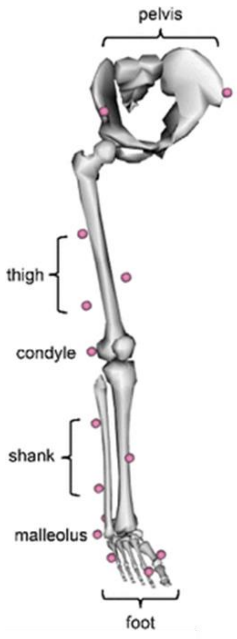


Figure 3.9: Examples of typical marker locations on the lower limb. Adapted from (Slater et al., 2018)

Model Calibration

Once a marker set has been defined, it is necessary to ensure its deployable within the SP software by creating a model template that includes information on the marker set, a distinct label for each marker and distinguishes the different segments and their possible relationship.

The primary reason for creating this model is to allow for the anatomical scaling of the marker set to the participant through a model calibration. By scaling the marker set the virtual markers described above can be estimated to define the LCS of a segment. In addition, this scaling technique can also be used to assist in auto-labelling of the markers in the processing of the subsequent movement trials, which is a common feature in SP software. Model calibrations consist of a short (approximately 10 second) static pose of the participant standing once all necessary markers are attached. Once the pose has been recorded via the SP system, the model template is applied, and a scaled version of the model to the subject is created via internal processing pipelines in the SP software (Systems, 2023). This subject specific model is then used for all other trials collected.

A possible extension of the model calibration is the addition of a second calibration that incorporates dynamic movements of the participant. This additional step can either be used to improve the auto-labelling for the more complex dynamic movements expected in the movement trials (i.e., incorporating a sit to stand or lunge), or can be used as a method to functionally calibrate the virtual markers of the LCS, such as that at the hip, by dynamically moving the joint through its range of motion (Piazza et al., 2004). This additional calibration is beneficial in further improving the subject specific model already created.

Multicentric Study

Within the context of the multicentric study, the assessment and subsequent outputs from the SP system data collection will focus on a high-level accurate estimation of the spatiotemporal features as a reference outcome for validation of the single sensor solution. As such, the marker set must be tailored to allow the quantification of the basic temporal (e.g., gait events) and spatial (e.g., step length) features of gait. In addition, the application of this marker set across patient cohorts and laboratories requires that the marker set must be easy to wear and will induce minimal variations between operators. Furthermore, the use of different SP systems and manufacturers across laboratories will mean that the model template created for the software and steps for the model calibration must be tailored to be compatible with the specification of each software.

Aim

Taking this into account, the remainder of this section aims to define a bespoke marker set that can be used to quantify the basic temporal and spatial features of gait that are easy to compare to the other sensor systems used in the lab-based protocol. To facilitate this, the marker set, SP model and calibration must be standardised across sites and SP manufacturer software and the level of possible error in application of the marker set must be minimised to ensure comparability across sites and participants. The rest of this section will describe the

bespoke marker set, the additional tools developed and requirements for its application across sites and SP manufacturer software.

3.3.2 Methods

Marker Set

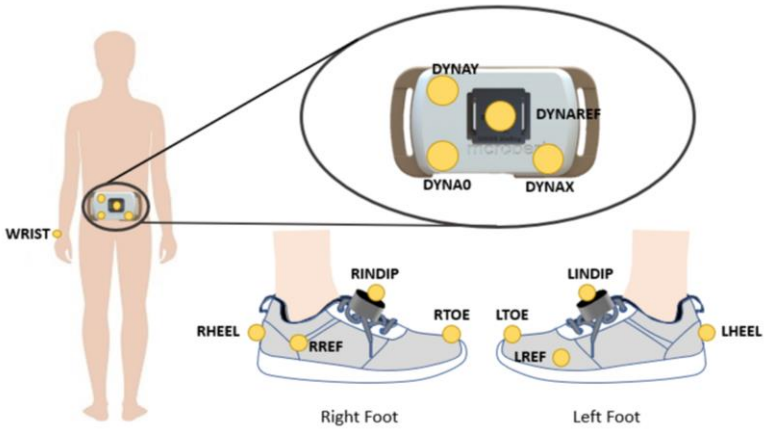
As the primary use of the SP data collection was the estimation of spatiotemporal parameters the main requirement of the marker set was precise tracking of the gait events (both ICs and FCs). Based on the requirements from common algorithms used in calculating gait events with an SP system, markers would be needed on the heel and toe of each foot. An additional requirement was to maximise the comparability of the SP data collected to the other measuring systems used in this data collection, which included the single sensor solution attached to the lower back and a multi-sensor system that included four IMUs; with one attached to the dorsal aspect of each foot, one attached to the non-dominant wrist and the fourth to ideally be situated in the same position as the single sensor solution, with LCSs of the two sensors aligned. As such, additional tracking markers would be attached to the top of the IMUs on the feet and wrist, and an LCS on the lower back that corresponded to the other two sensors would need to be constructed with markers.

To create a fixed structure that could ensure alignment of the LCSs of each sensor and the SP markers on the lower back, while also ensuring minimal variability of this setup across participants and sites, a bespoke 3D printed casing was designed. The case was designed to encapsulate the two sensors, ensuring correct alignment of the corresponding LCS. Additionally, three SP markers were adhered to the outing of the case in precise locations, based on divots created in the casing, to replicate the LCS in the SP data collection. As this structure was designed to not be opened unless necessary, access to the ports of both sensors was incorporated to allow easy charging and data transfer (Figure 3.10). To maximise comfort for the participants, all edges were bevelled, and the corners of the case rounded.

Figure 3.10: 3D printed case designed for the lower back.

With the current requirements resulting in only three markers being present on each segment, it was decided, as a safeguard for marker tracking and possible pre-processing, that an additional tracking marker would be attached to each foot and the case on the lower back. As the marker set would be minimal when compared to the typical marker sets used in human movement, the addition of a tracking marker on each foot could help distinguish right and left by placing the markers asymmetrically (i.e., different points on the lateral side of the foot). As the wrist sensor and tracking marker were not to be used in the calculation of any outputs from this study, a single marker was deemed appropriate.

Based on the requirements described above, a bespoke marker set was defined (Figure 3.11) and each marker given a distinct label.



Segment in SP Marker Set	Marker Label	Location
Left Foot	LTOE	Dorsal aspect of 2nd metatarsal head
	LINDIP	Top of Left Foot IMU
	LREF	Lateral aspect of 5th metatarsal head
	LHEEL	Calcaneus (approx. same height as LTOE)
Right Foot	RTOE	Dorsal aspect of 2nd metatarsal head
	RINDIP	Top of Right Foot IMU
	RREF	Lateral aspect of shoe, below the lateral malleolus and approx. the same height as RTOE
	RHEEL	Calcaneus (approx. same height as RTOE)
Lower Back	DYNAREF	Top of Lower Back IMU
	DYNA0	Fixed to lower back case
	DYNAX	Fixed to lower back case
	DYNAY	Fixed to lower back case
Wrist	WRIST	Top of Wrist IMU

Figure 3.11. An illustration of the marker configuration

Model Calibration

The main requirements when creating the model template and defining the calibration procedures, was to ensure ease of deployment across different SP manufacturer software, minimise operator burden when running the calibrations as well as optimisation of the procedures to improve the auto-labelling and use of the other sensor-based systems in the data collection.

To deploy the model template across systems, the specifications required in the template for each system was considered. As two different manufacturers of SP systems were included in this study, a template would need to be made for each, following the file compatibility and format required by each manufacturer.

Although the pipeline for processing the static model calibration varied between the SP system manufacturers, the general requirements of the trial (a short static pose) remained the same. Each laboratory was advised to perform this part of model calibration based on the SP manufacturers protocols.

As a dynamic calibration was required for the multi-sensor system also used in this data collection, it was decided to utilise this to include a dynamic trial in the model calibration. This trial included a range of movements, that were either necessary in calibration of the multi sensor system or would be regularly performed in the movement trials. The result was a dynamic calibration that consisted of multiple small tasks, with the participant being asked to start in a static standing position and lift and hold each leg, one after the other, for approximately 5 seconds, then raise the arm with the sensor attached above their head and back down before performing a short walking task that concluded the dynamic trial (Figure 3.12). To ensure inclusivity across participants, an option that included sitting for the first portion was used as an alternative to standing on one leg for an extended period.

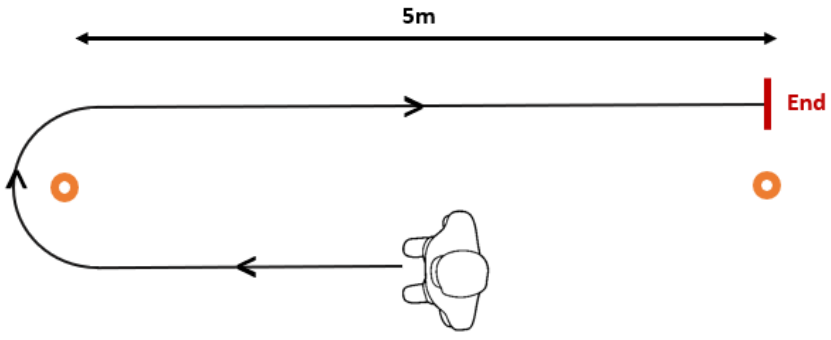


Figure 3.12 Schematic of the walking trial performed during the dynamic model calibration.

3.3.3 Conclusion

The aim of this methodology was to define and deploy a bespoke marker set and necessary calibration procedures for further processing in SP software that allowed easy comparison of the basic temporal and spatial gait features to the other sensor systems used in the lab-based protocol. A crucial aspect in successfully completing this aim was the standardisation of marker set and calibration procedures across SP manufacturer and sites. This was demonstrated well in deployment of the marker set, and model calibration procedures in the multicentric data collection, with all laboratories and operators correctly applying the markers and calibration for all 108 participants involved in the study.

The definition of a novel a reduced marker set, compared to the traditional marker sets used in the assessment of gait, aided a short data collection setup and good tracking of the markers during the trials, which subsequently improved time efficiency in a protocol that was already time intensive with the use of two sensor systems in addition to the SP system. Most importantly, the choice in marker set proved to be suitable for estimating the gait outcomes of interest as shown in Bonci et al (2022) and was efficient in speeding up the pre-processing of the SP data, as described in detail in the next section.

3.4 After the Data Collection: Standardising Data Pre-Processing

3.4.1 Background

Pre-processing of SP data is critical prior to calculating the desired output parameters, regardless of the context. A primary factor that needs to be considered in the pre-processing of SP data is the management of possible gaps in the recorded marker trajectories. As mentioned previously in this chapter, the reconstruction of the SP markers 3D coordinates is dependent on it being seen simultaneously by at least two cameras, with the addition of more cameras improving the marker tracking precision and possibly safeguarding from any obstructions to a cameras view, such as arms swinging during gait, the use of walking aids or subject rotation (Chiari et al., 2005). However, in the eventuality that a marker is not seen by two cameras, there will be a gap in the trajectory until it is seen by two or more cameras again (Figure 3.13). Although the operator can try to minimise the possibility of experiencing gaps in the marker trajectories during the system setup (e.g., by ensuring adequate positioning of cameras) it is an inevitable issue in complex movements or data captures close to the edges of the capture volume, both of which are apparent in the multicentric study.

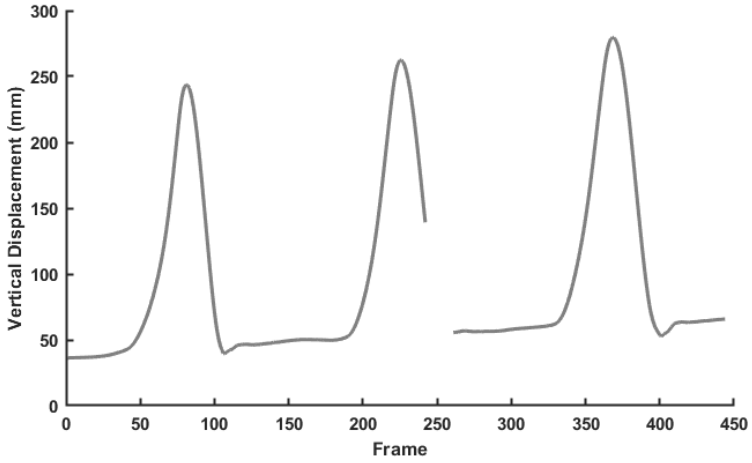


Figure 3.13 Example of a marker trajectory with a gap due to occlusion

A common approach to manage the occurrence of gaps is to "gap fill" using an appropriate method of interpolation. Gap filling of SP data is typically performed manually in the proprietary software of the system and requires the operator to visually inspect each gap before subjectively deciding which of the manufacturers interpolation methods to use (Vicon Motion Systems, 2023), while ensuring the data is not over manipulated due to the interpolation method chosen, which commonly worsens the bigger the gap (Figure 3.14). Depending on the number, length, and complexity of the recorded tests, as well as the

marker set used, this exercise can become extremely time intensive and burdensome. Moreover, as the most common methods of interpolation work on the assumption that the marker trajectory is present before and after the gap, and that the markers (including markers on the same segment which may assist in the interpolation) are accurately labelled (Howarth & Callaghan, 2010), the manual input from the operator could proliferate if the automatic labelling of the markers fails. As this whole process relies heavily on manual input and subjective decision making, the quality of pre-processed data can be influenced by the experience of the operator and subsequently question its position as a gold standard measure.

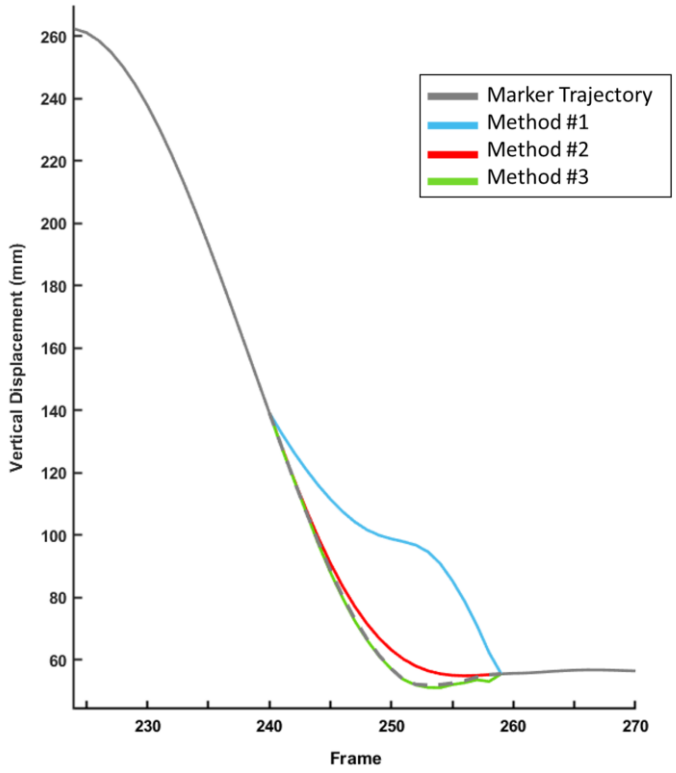


Figure 3.14 Examples of three different interpolation methods (blue, red and green) applied to a gap in the markers trajectory, (full grey line is the trajectory before and after the gap and the grey dahsed line showing what the trajecotry would be without the gap).

Variation in the quality of pre-processed data is also affected by which SP manufacturer software is used, with each manufacturer offering a different group of interpolation methods and tailored thresholds for their implementation within the software (Qualysis (Qualysis AB, Gothernburg, Sweden) or Vicon (Oxford Metric, Limited, Oxford, England)). For the comparability of data that has variation in operators and SP manufacturers, as is the case in the multicentric study, an automatic pre-processing workflow that is SP system agnostic would be required. Although previous literature has considered the development

of automatic methods, by using machine learning to recover missing markers (Liu & Mcmillan, 2006) or data from additional sensors to estimate marker position (Bobilev et al., 2012), these require an additional set of "clean" data to train the model or additional sensors attached to the participants, both of which are not feasible in the multicentric study.

A possible alternative approach is the development of a gap-filling toolbox, that can be used to pre-process SP data externally from the proprietary software. This would still require manual review of the output data, but variation caused by the differences in software between manufacturers would be minimal.

As the data in this study will include a standard marker set (defined in Section 3.3) and task protocol, the use of a gap-filling toolbox could be optimised into a bespoke pipeline that includes specific criteria (i.e., set a maximum length for gaps suitable for interpolation) and a hierarchal approach to prioritise certain interpolation methods suited to the data. Furthermore, by compiling the pipeline into a pre-processing workflow, additional features such as a quality check of the marker's labels and identification of any mislabelling could assist in automating some aspects of the manual work required prior to gap filling.

The aim of this work is to develop a semi-automatic workflow integrated in a graphical user interface (GUI), to implement quality checks and the necessary procedure of gap filling for the pre-processing of SP data. The deployment of this GUI in the multicentric study is intended to ensure the comparability of SP data pre-processed by different operators and collected with different SP systems, to maintain the gold standard status of the measure.

3.4.2 GUI Integration

The primary reason for choosing a GUI format, is the ability to compile functions in a workflow that includes breaks, to allow for manual review or amendments before progressing to the next stage. Additionally, the development of a user-friendly interface would allow all operators, regardless of their previous experience working with SP data or the programming language, to complete the pre-processing (Figure 3.15). Development of the source code and compiling of the GUI were completed in MATLAB 2020a (MathWorks, Natick – USA).

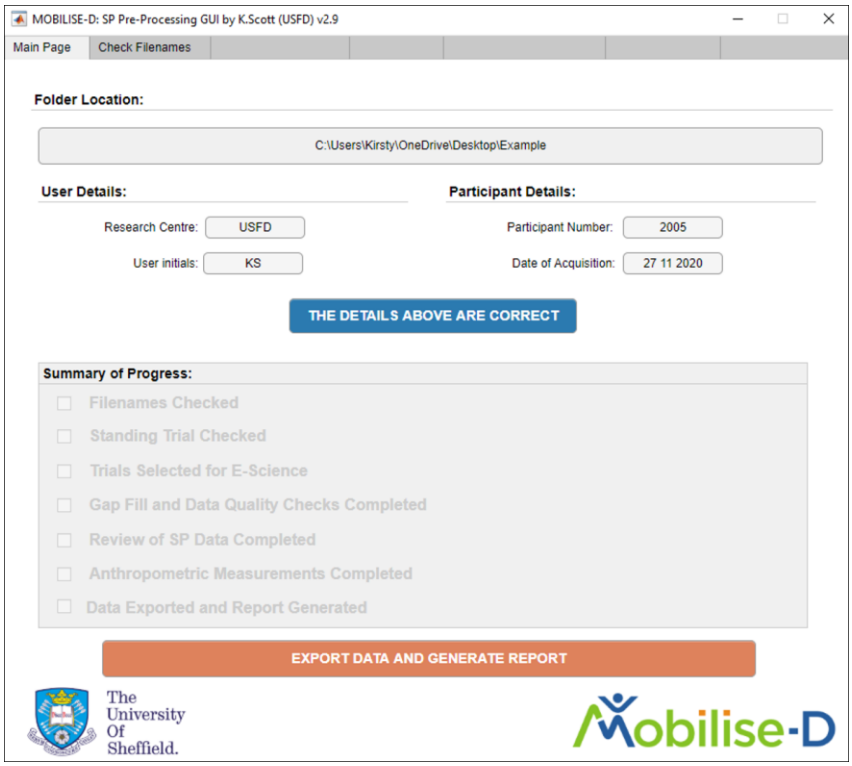


Figure 3.15: Main page of GUI.

Prior to the development of the GUI, it was necessary to consider the steps required in the pre-processing of SP data, and determine which parts could be automated and standardised, to guarantee the level of data quality expected from a gold standard measure. As a general rule, good data quality is characterised by its accuracy, completeness, and consistency (European Medicines (Agency, 2010). In the case of SP data, the primary factors that can compromise the quality are the mislabelling of markers, over manipulation of marker trajectories when implementing gap filling methods, and lack of standardised procedures to assess the quality and comparability of pre-processed data across datasets. Based on the factors identified, a semi-automatic workflow, that contained four main stages, was proposed (Figure 3.16):

Setup and Initial Checks: This initial stage is used to input specific variables linked to the dataset into the GUI. A quality check for the filenames of the SP data is completed, to ensure all filenames fit the format defined in the study. Finally, an initial check of marker labelling in all trials is completed and reference information of the geometrical relationship of markers on the same segment is generated from the static pose used in the model calibration.

Automatic Pre-Processing: A bespoke pipeline that includes detection and amendments to mislabelled markers as well as implementation of the interpolation methods chosen for gap filling is ran on all task trials included in the data collection.

Trajectory Screening: For quality assurance of the automated pre-processing stage above, a manual check of the correctly labelled and filled markers trajectories is essential to determine if any manual amendments or final stages of pre-processing are required.

Final Check and Export: A final check to determine the quality of the pre-processed data is executed and a standardised report is generated to allow transparency of these results across datasets. Finally, a standardised csv format of each SP trial is exported.

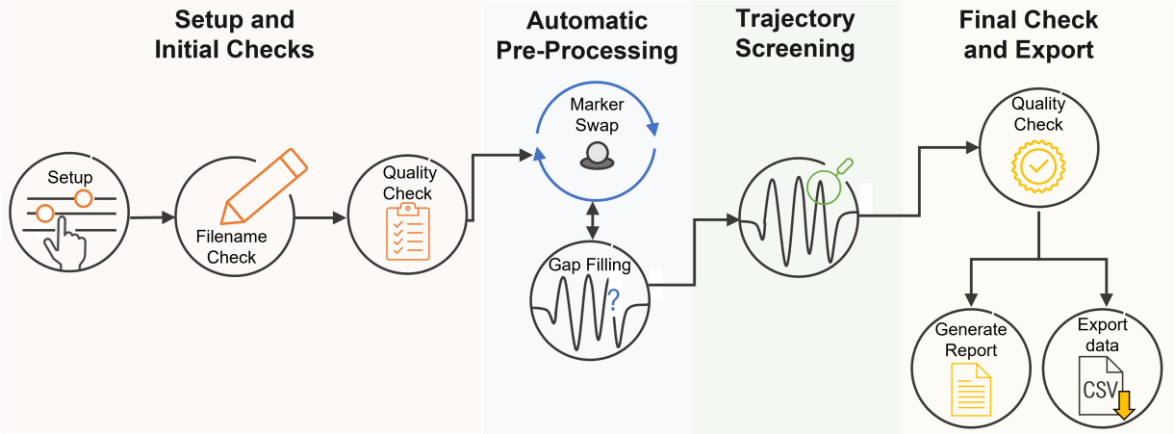


Figure 3.16: Schematic of the proposed pre-processing workflow divided into the four stages.

As it was deemed appropriate to include a manual check of the data output from the automatic pre-processing pipeline, it was crucial that the operator could access the SP data easily between the GUI and proprietary SP software during the workflow, with any amendments to data made in a file format that was readable and editable in both.

To fulfil this requirement, a c3d file format was chosen as the input file for the GUI, due to its universal use by different SP software. All data processing in the GUI was amended to the c3d file, allowing an easy switch between software prior to the export of the final data in csv format. Reading and amendment of the c3d files in the GUI was supported by the Biomechanics Toolkit (Barre & Armand, 2014) and utilised in the development of the GUIs functions.

To further improve the switch between software and accommodate pauses in the pre-processing of the data, memory files were embedded in the GUI and generated in a subfolder in the SP data directory, once the initial setup was complete. This allowed users to close and reopen the software, with memory of the previous data used and location in the workflow automatically set in the interface, ensuring no repetition of previous steps was required.

3.4.3 Development of GUI functions

Development of the GUI functions was split into the four stages of the pre-processing workflow defined in Figure 3.16. Initial development of the functions and related algorithms was based on pilot data of the lab-based protocol collected at each of the five laboratories used in the multicentric study (please see Appendix Three for internal ethics approval). This data was used as a census check to ensure the defined thresholds within certain algorithms were appropriate for a smooth integration within the GUI.

3.4.3.1 Stage 1: Setup and Initial Checks

As mentioned above, a feature included to improve the GUIs usability and accommodate breaks in the pre-processing of data, was the creation of a memory file (txt.) that included specific information for each dataset. To input the information needed in this file, the user would first carry out a short setup of the GUI. This included navigation to the SP data folder directory, input of the lab the data was collected, the participant ID (created specifically for the study) and the date of the acquisition. As an additional check, the automatic function reading the SP directory would also confirm that c3d files were present in the folder and alert the user if none were present. As it was crucial that the filenames were correct for each task for the data standardisation with the other sensor systems used, filenames of the available c3d files were reviewed, using information in the newly generated memory file, and any errors between the actual and expected filenames were highlighted in yellow via an internal report generated by the GUI (Figure 3.17). This was also used as an opportunity for the user to review the available c3d files to ensure all expected trials were present.

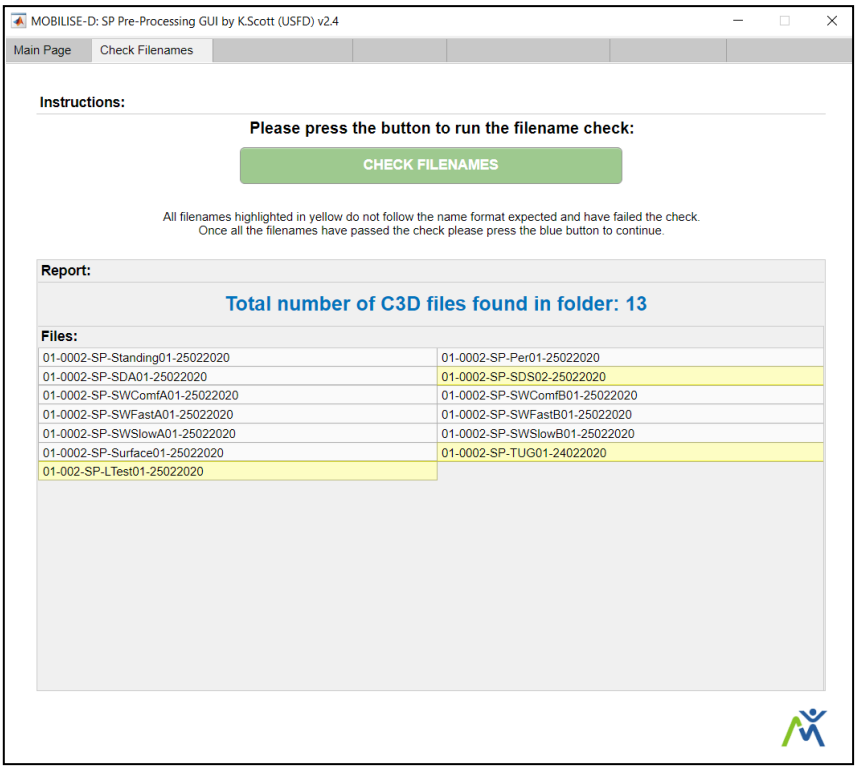


Figure 3.17 Example of the internal GUI report for filename check, which indicates the total number of files, each files name and highlights in yellow if the filename does not meet the expected format.

Once complete and all amendments by the user were made where needed, an initial quality check of the marker set was executed. First, the c3d file for the static trial used in the model calibration was read and the marker labels were reviewed to confirm that all 13 labels were present (markers in Figure 3.18 and the additional wrist marker). To ensure that each marker was labelled correctly in each segment (i.e., the "*RHEEL*" label was the marker on the right heel) a heuristic assessment of the marker's linear geometrical relationship to markers on the same segment (*Lower Back*, *Left Foot*, and *Right Foot*) was applied (Figure 3.18).

For the *lower back* this assessment considered the marker with the greatest and second greatest vertical displacement to the ground as the "*DYNAY*" and "*DYNAREF"* respectively. The other two marker labels were then determined based on the inter-marker distance to 'DYNAY', with the greatest distance deemed as the "*DYNAX*" marker.

For the *Right* and *Left Foot* this assessment considered the marker with the greatest vertical displacement to the ground as "*R/LINDIP"*. The "R/LTOE" and "*R/LHEEL"* markers were identified as the having the greatest inter-marker distance and then individually assessed based on the inter-marker distance to the "*R/LREF"*. To safeguard for possible error in

marker position of the "*R/LREF*" marker, the distance between the heel and toe markers between segments was also considered.

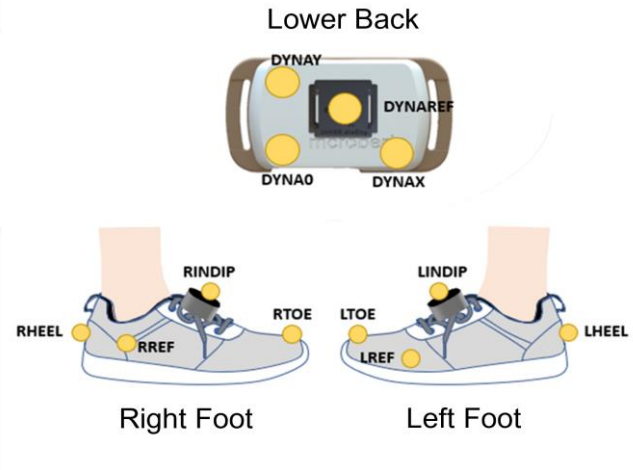


Figure 3.18 A schematic of the marker set, which was used to determine the criteria for the heuristic assessment of the markers labelling.

If passed, reference information of the markers XYZ position and geometrical relationship to the other markers on the same segment was exported and saved to later be used in the automatic pre-processing stage. If any errors were flagged, the user was alerted via the GUI and asked to review and amend labelling via the SP software.

To improve the automatic running of the pre-processing stage, a short check to ensure all 13 markers were present in the task trials was completed and each trial was automatically assessed for the presence of gaps by frame (sampling frequency of 100Hz, with 0.01 seconds = 1 frame) and for each marker (e.g., a 13xN logical array, where N is the number of frames and 0 = no gap and 1 = a gap). This information would later be used in the pre-processing stage and allow comparison between the raw and pre-processed data, to determine the efficiency of the pre-processing.

3.4.3.2 Stage 2: Automatic Pre-Processing

To facilitate any troubleshooting that may be necessary from errors in the pre-processing stage, a backup of the raw c3d file for each trial was automatically created prior to running any pre-processing. Following this, a bespoke pre-processing pipeline was executed for each trial. This bespoke pipeline included four main algorithms to accommodate its execution. Firstly, a check to ensure there was no mislabelling in the markers at the start of the trial, which would subsequently lead to mislabelling throughout the rest of the trial. Second, an assessment and amendments to mislabelling within the trial. Third,

implementation of a rigid body interpolation and finally, implementation of a pattern-based interpolation. It is worth stating that no automatic pre-processing was applied to the single marker on the wrist as this was not deemed to be necessary for the primary output of multicentric data collection.

Mislabelling in first frame

Although algorithms for automatic labelling of markers are commonly integrated within SP software, their execution can give varied results. Errors in the execution of automatic labelling is likely at points in the trial where tracking of markers is reduced due to insufficient coverage from the cameras (e.g., at the edge of the 3D capture space or occlusion from equipment, such as chairs, being incorporated in the recorded task). If this error in labelling occurs within the first frame of recording, the error will propagate throughout the trial and deem any automatic pre-processing invalid. As the multi-centric task protocol would include starting points at the edge of the 3D capture and starting positions seated in a chair, it was crucial to identify this possible occurrence of mislabelling before proceeding with the rest of the pre-processing.

As the skeletal structures of the body are considered as simplified rigid segments (as described in Section 3.3), it could be assumed that at any point of recording the inter-marker geometrical relationship on a same segment would remain constant. Using this logic, the reference information of the inter-marker geometrical relationship (exported from the static trial used in the model calibration) (y) could be compared to the inter-marker geometrical relationship ($\hat{y}$) for each frame (i) of each task trial:

$$Difference_i = (y_i - \hat{y}_i)$$

However, this is an oversimplification due to the accumulation of error in marker reconstruction (described in Section 3.2) and the additional error from markers attached directly to the body due to the morphology to skin during activity, and in the case of the foot, changes in the shape of the skeletal structure during movement (Chan & Rudins, 1994).

In light of this consideration, this approach could still be used at the start of each task trial due to a period of static pose, prior to the participant engaging in the task at hand but would require the inclusion of two additional factors; the change in geometry due to the random and systematic errors from the marker reconstruction (as detailed in Section 3.2) and possible morphology to the foot markers. By creating a window of deviation that accounts for both factors, a threshold could be set to determine if a segment was mislabelled from the beginning of the trial.

For application of this method, windows of deviation for the error in marker reconstruction and morphology to the feet must be defined. Based on initial findings of the maximum errors associated to marker reconstruction (found in Section 3.2) and minimum inter-marker distance expected across segments being 50mm (from markers on the *Lower Back*), a deviation window of ±10mm (SD_m) was deemed a safe threshold. To accommodate morphing of marker position on the foot, an additional window of ±40mm (SD_f) was defined as a conservative deviation, due to the minimum inter-marker distance on this segment from the pilot data being ~75mm. Based on these defined thresholds, $Difference_i$ could now be determined if this was mislabelling of the segment when $i = 1$.

For the lower back segment this would be considered a mislabelling if:

$$\boldsymbol{Difference_i} < \left(\boldsymbol{y_i} - \boldsymbol{SD_m}\right) \textbf{ OR } \boldsymbol{Difference_i} > \left(\boldsymbol{y_i} + \boldsymbol{SD_m}\right)$$

For the foot segments this would be considered a mislabelling if:

$$\boldsymbol{Difference_i} < \left(\boldsymbol{y_i} - (\boldsymbol{SD_m} + \boldsymbol{SD_f})\right) \textbf{ OR } \boldsymbol{Difference_i} > \left(\boldsymbol{y_i} + (\boldsymbol{SD_m} + \boldsymbol{SD_f})\right)$$

If mislabelling of a segment was identified, the user would be alerted to this trial via the SP software, prior to continuing with the pre-processing pipeline.

In the instance that a marker on the segment was not present in the first frame, the function would progress to the next frame, until all markers were present, and base the assumption of mislabelling on this frame instead. As a precaution to this method taking a point after the initial static pose, a velocity threshold was set on the other markers present (Bonci et al., 2022). In the case that the function was unable to determine mislabelling (i.e., there was no portion of the initial static pose where all markers on the segment were present), a flag was activated to not proceed with the pre-processing of this segment in the selected trial. At the end of running the pre-processing pipeline, the user was then notified of this flag via the GUI and asked to manually inspect for mislabelling in the trial prior to re-running the pre-processing pipeline for this specific trial (Figure 3.19).

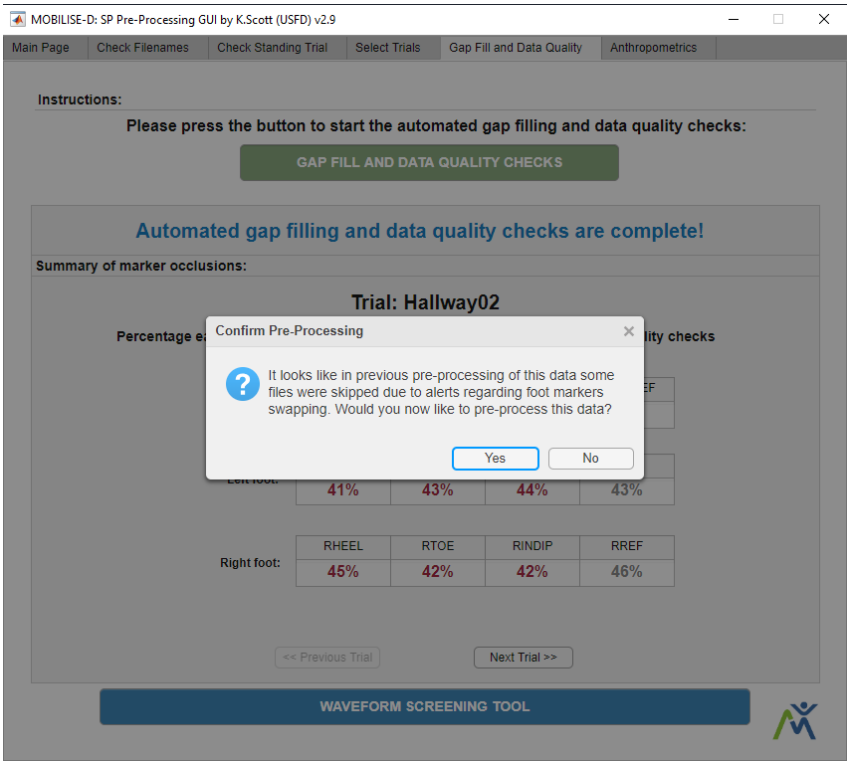


Figure 3.19 Example of the GUI flag to inform the user of mislabelling errors.

Mislabelling within the trial

As gaps are directly linked to occlusion of the marker to the cameras view, it is common for a marker to be mislabelled just prior and after a gap due to a decrease in cameras tracking the marker. If this mislabelling is not corrected, subsequent gap-filling will lead to extreme deviations in the marker's trajectories (Figure 3.20). To combat this error, a function that would be efficient in flagging and amending any mislabelling within the trial would be required.

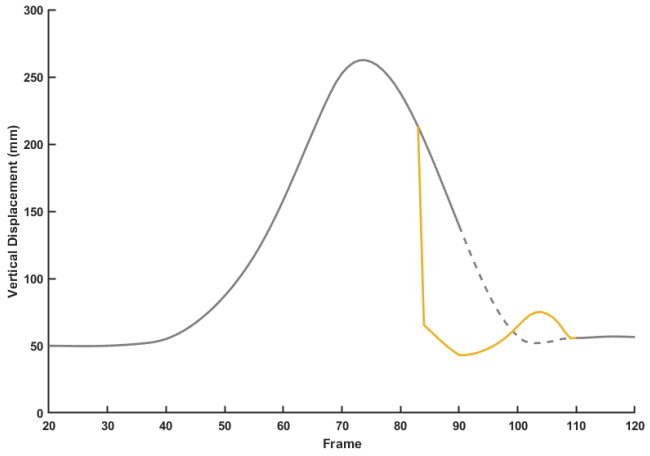


Figure 3.20 Example of gap filling (yellow line) when marker is mislabelled prior to gap compared to actual marker trajectory if no gap was present (grey dashed line)

During the initial development of this algorithm, difficulty arose in including the foot segments within the timeframe of its deployment due to the extreme changes in marker relationships during walking. This decision was, in turn, backed by the pilot data used in the algorithm development, that determined mislabelling within the trial was rarely seen in the foot segments, when compared to the lower back segment, due to camera setup in the laboratories primarily focussing on the feet, larger distances between the markers and asymmetry in the marker setup between left and right. As such, this function would solely focus on correcting labelling for the *Lower Back*, with additional features included in the later stages of pre-processing to correct any mislabelling of the feet.

With the marker configuration on the *Lower Back* designed to be robust across participants and laboratories by adhering the markers to the case, the assumption that minimal changes would be found in the marker's geometrical relationship on this segment regardless of activity was found to be valid. As such, a more advanced heuristic assessment of the inter-marker geometrical relationship (that included assumptions on the linear and angular measures) was developed from the assessment used in Stage 1 for the static trial.

As this assessment would be based on the relationship of the markers, amendments to mislabelling if only one marker was present would not be possible. In addition, during pilot testing of the algorithm, the robustness in amending mislabelling if only two markers were present was not satisfactory to allow confident implementation of the successive gap-filling algorithms. As such, a minimum of three markers on the segment would need to be available.

Using the logical array of gaps in the marker's trajectory generated in Stage 1, gaps were identified and a sliding window of 1 frame (0.01 seconds) was applied to find the last point before the gap and the first point after the gap where at least three markers on the segment were present. Once identified, the difference in the geometrical relationship of the markers on the segment were compared to the reference data and mislabelling was identified as per the description in the *Mislabelling in first frame* algorithm. If mislabelling was identified, when compared to the window of deviation threshold SD_m (i.e., ±10mm), a sequence of assumptions based on the linear and angular relationship of each marker to the others in the segment were executed. Based on the output, the markers were then relabelled to meet the geometrical criteria. This process was modified to run whether three or all four markers were present. Once complete for this frame, the sliding window would then move to the sequential frame (either the moving backwards for before the gap, or forwards for after the gap) and the same process of checking for mislabelling and amending would be complete until the first and last frame on either side of the gap that contained all four markers labelled correctly was found, or there was the instance of another gap. At the

end of running this function the modified marker trajectories were amended to the c3d file and saved to allow use of the new trajectories in the implementation of the gap filling algorithms and manual review in the SP software.

Interpolation Algorithms

To ensure a high level of confidence in the automatic interpolation of gaps, two methods were selected as appropriate, based on the pilot data collected and the expertise of the author. These were a rigid body interpolation and a pattern-based method, which both use "donor" markers (i.e., markers on the same segment that do not have a gap) as input to the interpolation.

To deter from any over manipulation of the foot marker trajectories, which may elicit errors in the calculation of the gait events, it was determined that a maximum length of gap to interpolate should be defined. Based on the average gait event duration being 0.5-0.6 seconds (Murray et al., 1964), a maximum gap length of 50 frames (0.5 seconds) was set. Due to the rigidity of the case on the *Lower Back* and at least three markers on this segment being necessary to calculate the segments specific outputs, a maximum gap of 250 frames (2.5 seconds) was set.

The algorithms described below have been previously defined within the literature (Camargo et al., 2020) however, for their implementation within the GUI, the author developed the programming script and optimised for their implementation within the GUI, as per the requirements defined above.

Rigid Body Interpolation: The rigid body fill requires three markers on the segment to be present before and after the gap. The function utilises the Kabsch algorithm to determine the optimal rotation matrix, whose rows are the position vectors of the three donor markers from the segment, between the last frame prior to the gap ($i = 0$) and each frame (g) of the gap. Based on the offset between the rotation and translation of the donor markers, interpolation of missing marker is completed. As this method of interpolation works on each frame of the gap individually, its implementation can vary between frames, with frames with less than three donor markers being rejected.

Pattern based interpolation: This algorithm takes the offset from one donor marker between its actual data and a linear interpolation and adds that offset to the linear interpolation of the desired marker. If there are two valid donor markers, the one with closest average position to the desired marker at the frames before and after the gap will be selected.

To remedy any interpolation in noisy portions of a trial, a quality check that again compared the geometrical relationship was made, however, in this instance reference information was taken from the last prior to the gap where all markers were present and all segments used the defined window of deviation threshold of SD_m (i.e., ±10mm). If passed, then the interpolated trajectory was amended to the c3d file. If the check failed, the interpolation was rejected and the information on why the gap was not filled was logged for the operator to review in the next stage.

Development of pipeline

Using the pilot data collected to develop the algorithms, the pipeline order was optimised (Figure 3.21) to improve the efficiency of pre-processing for the multicentric specific data, while still considering the processing time of the pipeline's execution.

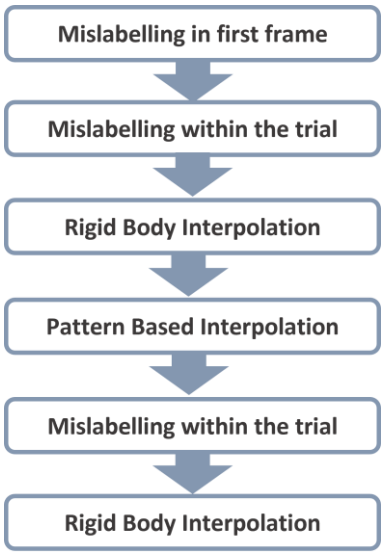


Figure 3.21 Final order of automatice pre-processing pipeline.

3.4.3.3 Stage 3: Trajectory Screening

As a manual check was deemed a necessary and smart approach as quality assurance of the automated pre-processing, a screening tool that displayed each markers trajectory was developed. This tool graphed the 3D marker trajectories in the XYZ direction, with separate for each segment to reduce the number of pages checked and aid in identifying any anomalies. To further assist in reducing time of the manual check, two patch shades (red and grey) were created to define the reason for a gap still existing in the trajectory, indicating whether manual gap-filling via the SP software was deemed appropriate (red) or not (grey) (Figure 3.22).

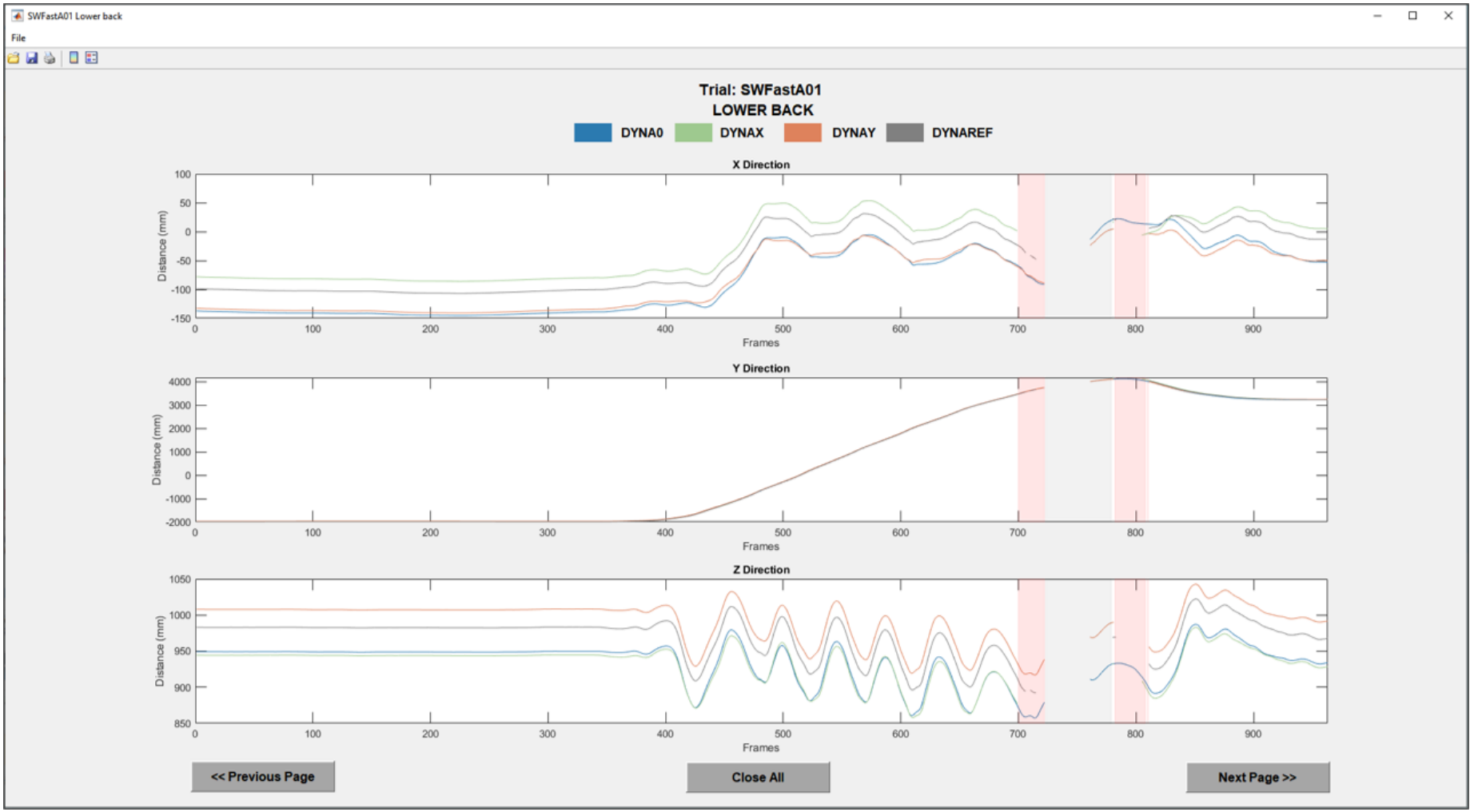


Figure 3.22 An example page from the trajectory screening tool showing the marker trajectories for the lower back segment. The grey patch indicating a gap that cannot be filled and the red indicating that manual review and possible interpolation is recommended.

The criteria for each shade were defined as:

- *Grey patch (gap is not suitable for manual amendments):*
    1. The gap occurred right at the start or end of the trial, meaning no data was available for the interpolation.
    2. The gap is longer than the maximum length defined for interpolation (e.g., >50 frames for the feet and >250 for the lower back)
- *Red patch (recommendation to manually review and, if appropriate, gap fill):*
    1. There are not enough labelled markers to automatic check mislabelling of interpolate gap.
    2. The automatic interpolation did not pass the quality checks. This primarily happened when the marker data is very noisy.

To improve the usability of this tool, the graphs appeared in order of trial, with active pushbuttons allowing the user to navigate to graphs of the segments, giving them full control to review graphs multiple times and in their preferred order. In addition, this graphing tool was accessible to run simultaneously with the SP software, allowing the graphs to guide a more time efficient manual check in the software.

To promote the use of this manual review, users were required to manual check that it had been completed via the GUI before being able to progress to the final stage.

3.4.3.4 Stage 4: Final Checks and Export

Finally, final quality check was completed on the data, by determining the reduction in the number of gaps from the 'raw' to the automatically pre-processed data. This information was exported as a 'data quality check' pdf. report, for the user and author of the GUI to review for final authorisation that the data met the standards necessary for the future processing and standardisation within the multicentric studies dataset (Palmerini et al., 2023). The c3d data files for each trial were converted into a csv format to meet the standardised file format of the multicentric study, before finally uploading the csv files and the 'data quality check' pdf report to the study platform.

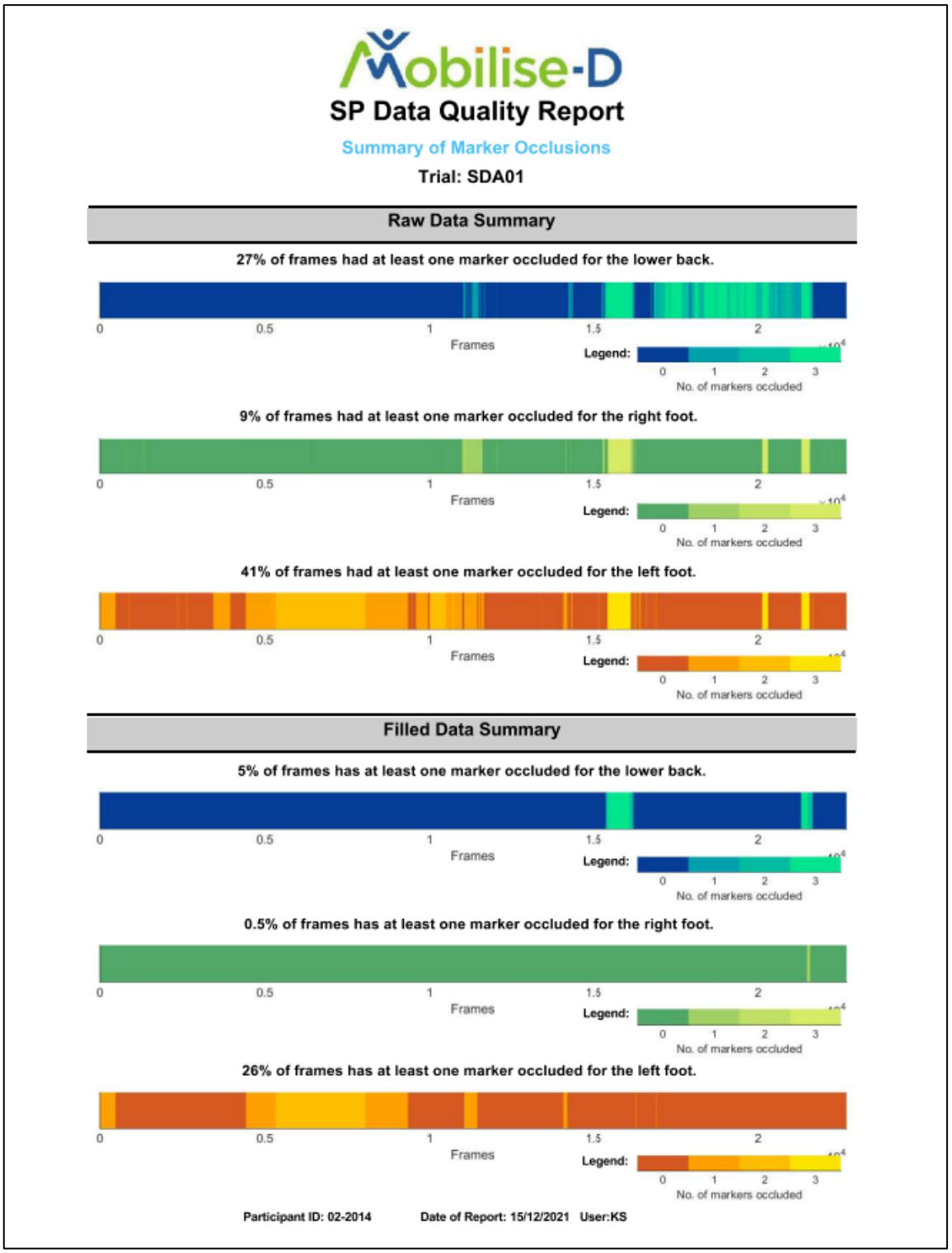


Figure 3.23 Example of the pdf data quality report. The first section visualises the gaps in the raw data, with the blue bar showing gaps in the lower back segment, the green bar showing gaps in the right foot and the orange showing gaps in the left foot segment. The lighter colours in each bar represent points in the trial with gaps, with the lighter the colour indicating the higher the number of makers with a gap. The bottom section highlights the same information for the data after the automatic pre-processing.

3.4.4 Deployment

To ensure deployment of the GUI was smooth across laboratories and users, a user guide was developed (please see Supplementary Material 'Pre-ProcessingGUI Manual.pdf') and a training session for each lab was organised by the author. Due to the varying experience of the users, a trialled approach that included the author checking the uploaded csv data via a modified version of the marker trajectory screening tool was completed for the first five datasets at each site. This trial method, aimed at ensuring the data quality met the requirements necessary for its use as the gold standard (data was only authorised for further processing after review from the author), while allowing the user to feedback any queries to the author and reduce the repetition of user errors. In general, the mix of open resources and one-on-one feedback proved to be successful, with all 108 datasets successfully being pre-processed via the GUI, regardless of the constraints to virtual training. In addition, the opportunity to feedback to the author allowed improvements in the GUIs usability, with updated versions of the software being released when deemed appropriate and subsequently leading to positive feedback of the GUIs usability from the users.

Regarding the efficiency of the automatic pre-processing pipeline, the pipeline reduced the occurrence of gaps from approximately 30% gaps in the raw data to <10% gaps in the pre-processed data, for each trial within the task protocol. The accuracy of the developed algorithms proved to be strong, based on the validation of the processed temporal outputs of this data from Bonci et al (2022), which found median absolute errors of <1% when compared to the gold standard pressure insoles that were part of the multi-sensor system also used in this data collection).

4 DESIGN AND VALIDATION OF A MULTI-TASK, MULTI-CONTEXT PROTOCOL FOR REAL-WORLD LIKE GAIT

A substantial part of the material presented in this section has been published in:

Scott, Kirsty, Bonci Tecla, Salis Francesca, Bertuletti Stefano, Caruso Marco, Alcock Lisa, Hansen Clint, Buckley Ellen, Schwickert Lars, Gazit Eran, Cereatti Andrea, Lorenzo Chiari, Sharrack Basil, Walter Maetzler, Becker Clemens, Hausdorff Jeffrey, Vogiatzis, Ioannis, Brown Philip, Del Din Silvia, Eskofier Björn, Helbostad Jorunn, Paraschiv-Ionescu Anisoara, Keogh Alison, Kirk Cameron, Kluge Felix, Micó-Amigo M. Encarna, Mueller Arne, Neatrour Isabel, Niessen Martjin, Palmerini Luca, Sillen Henrik, Singleton David, Ullrich Martin, Vereijken Beatrix, Yarnall Alison, Rochester Lynn and Mazzà Claudia. Design and validation of a multi-task, multi-context protocol for real-world gait simulation. J NeuroEngineering Rehabil 19, 141 (2022). doi: 10.1186/s12984-022-01116-1

Please see Appendix Two for author declarations.

4.1 Background

As discussed throughout this thesis, wearable devices such as inertial measurement units (IMUs) can be used to quantitatively assess both mobility qualifiers, with associated observations carried out in the form of short, structured tests (for example when instrumenting a six-minute walk test (Crapo et al., 2002) or a Timed Up and Go test (Podsiadlo & Richardson, 1991a) for mobility capacity or continuous unsupervised monitoring for mobility performance. In both cases, the signals from the IMUs are processed to extract specific features of mobility known as digital mobility outcomes (DMOs), such as the spatiotemporal parameters of gait. While algorithms to extract DMOs for capacity tests can be validated directly in a laboratory when equipped with a gold standard (e.g., 3D motion capture systems), the algorithm validation becomes much more complex when dealing with performance related DMOs. This is because measuring mobility in daily life entails dealing with confounding factors arising from multiple sources, including pathological characteristics, patient-specific walking strategies, environment/context, and purpose of the task. Accounting for all these factors within a single set of observations carried out within a limited laboratory space, while ensuring minimisation of participant burden and safety, is a very difficult and complex endeavour.

Previous work validating the estimation of DMOs have predominantly focused on standard straight walking assessments over a short distance (Aich et al., 2018; Han et al., 2019b;

Mico-Amigo et al., 2016) within a laboratory setting. Although this task is beneficial in gaining a measure under controlled conditions i.e., a measure of walking capacity, it does not consider any contextual factors or complexities in walking that that are a necessary aspect to include when validating performance DMOs. In response, recent laboratory-based validations have proposed protocols that included assessments with more complex tasks such as stair negotiation (Zhou et al., 2016), incline walking (Martinez-Hernandez & Dehghani-Sanij, 2018), inclusion of single and dual tasks (Dijkstra et al., 2008), variation in walking speed (WS) imposed via a treadmill (McGinnis et al., 2017) or overground (Lee & Park, 2011), as well as variation of walkway surfaces (Allet et al., 2008) and curvilinear paths (O'Brien et al., 2019). The fact that these studies mostly attempt to validate specific DMOs related to the task (e.g., validation of a cadence algorithm during stair ascent (Zhou et al., 2016) or detecting variation of temporal parameters on different walking surfaces (Allet et al., 2008)) limits the generalisability of their results and more complex scenarios mimicking the real-world have hence been proposed. These approaches attempt to simulate daily life environments for the purpose of validating continuous monitoring devices and usually entail the participant moving freely within a lab setting while completing a series of goal-oriented tasks designed to mimic the postures and movements expected to be seen in the real-world. In accordance with this concept, Bourke et al., (2017) identified a subset of tasks from the Compendium of Physical Activity (Ainsworth et al., 1993), to include as many variations as possible of real-world walking and associated postural transitions. This led to a protocol that had approximately 30 minutes of activity data for each participant, when deployed in a group of healthy older adults, and included 134 tasks with multiple repetitions of different transitions and straight walking at varying speeds. With such a comprehensive set of tasks however, this protocol is unsuitable for assessing individuals with reduced levels of mobility. Subsequently, protocols with a more refined list of tasks have been proposed by other authors to reduce repetitions of tasks (Hegde et al., 2018) or accommodate the inclusion of patients with Parkinson's Disease (Pham et al., 2017), but with these home-like assessments lasting 90-180 minutes the duration is still such that only a simple gold standard (such as 2D videos) could be used, limiting the validation of the monitoring device to activity recognition or basic gait parameters like step detection. Warmerdam et al., (2021) recently proposed a much shorter, and hence more feasible home-like assessment that was situated in the volume of a 3D motion capture system, but as the main aim of this set of tasks was to observe changes in balance and postural control the translation of this assessment when aiming to mimic real-world walking would not be effective. Overall, a comprehensive protocol that could effectively mimic a variety of complex walking patterns within a lab setting and be safely administered to participants with different levels of mobility has yet to be identified.

When attempting to move the validation of estimating DMOs to a real-world context, significant hurdles are associated with the feasibility of deploying a gold standard and ensuring a meaningful amount of data are collected. Although several technological solutions have recently been proposed in the literature, including GO-Pro body-worn cameras (Bourke et al., 2017; Hickey et al., 2016), footswitches or pressure insoles (Storm et al., 2016), foot or ankle-mounted IMUs (Paraschiv-Ionescu et al., 2019), and multi-sensor systems that integrate IMUs and infrared distance sensors (Bertuletti et al., 2019), the complexity of processing data from some of these systems often limits the duration of the observation. With no clear information in the literature about how long these real-world observations should last and which tasks must be included while still preserving an acceptable level regarding the burden and cognitive demand asked of the participant, an assessment of the duration required to ensure a meaningful amount of data are recorded while keeping in mind the above considerations is necessary. For example, if the attention is focused on assessing gait performance, data from an adequate number of walking bouts (WBs) should be collected and these should be recorded within complex and representative contexts, such as inclined walking, stair negotiation and indoor and outdoor settings.

Within this framework, the primary aim of this study is to validate a multi-task and multi-context protocol for simulating real-world gait. The protocol includes a laboratory-based assessment and a 2.5-hour unsupervised data collection in the participants' habitual environment. Validation of this protocol will focus on proving that the chosen series of complex activities in the laboratory-based assessment: a) are suitable for the evaluation of a variety of gait patterns, including healthy gait and impaired gait associated with neurodegeneration, a proximal femoral fracture, chronic pulmonary disease or congestive heart failure; b) include at least one WB, defined as a minimum of two consecutive strides of both feet (Kluge et al., 2021); c) induce a large variation in gait strategies, resulting in a broad range of WSs captured; and d) avoid redundancy in the tasks to minimise burden to the participant. In addition, a secondary aim of the validation is to determine whether 2.5 hour of unsupervised monitoring in the participants' habitual environment is a long enough observation to collect a set of data that is extensive and reliable for assessing the validity of gait related DMOs in a real-world context. We expect that these results will establish a common ground for the technical validity of wearable devices aimed at estimating gait related DMOs in real-world settings.

4.2 Methods

4.2.1 Data Collection

As part of multicentric study, a convenience sample of 108 participants were recruited from six cohort groups that included older healthy adults (HA) and participants with potentially altered mobility due to Parkinson's Disease (PD), Multiple Sclerosis (MS), Proximal Femoral Fracture (PFF), Chronic Obstructive Pulmonary Disease (COPD) or Congestive Heart Failure (CHF). These cohorts were chosen as presenting a variety of gait and mobility features specific to the pathology. Besides the cohort specific inclusion and exclusion criteria described in Table 4.1, all participants were: 1) able to give informed consent, 2) willing to wear the sensors setup and participate in the different data collections of the study, 3) scored >15 in the Montreal Cognitive Assessment (MoCA), 4) were able to walk four meters, 5) had no comorbidities impacting mobility or compliance. Data were collected across five gait laboratories after receiving written informed consent (Ethics approvals: The Newcastle upon Tyne Hospitals NHS Foundation Trust and Sheffield Teaching Hospitals NHS Foundation Trust: London – Bloomsbury Research Ethics committee, 19/LO/1507; Tel Aviv Sourasky Medical Center: the Helsinki Committee, 0551-19TLV; Robert Bosch Foundation for Medical Research: medical faculty of the University of Tübingen, 647/2019BO2; University of Kiel: medical faculty of Kiel University, D540/19). Participant demographics were collected, and patient characterisation was completed based on clinical assessments specific to each cohort (Mazzà et al., 2021) (Table 4.2).

Table 4.1: Inclusion and exclusion criteria adopted for the different cohort groups

Group	Inclusion criteria	Exclusion criteria
OHA	65+ years of age	
PD	-aged 18+ years -Diagnosis of PD according to the Movement Disorders Society criteria	-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	-impaired mobility related to non-PFF causes, as judged by the investigator

	diagnosis S72.0, S72.1, S72.2) as diagnosed on X-rays of the hip and pelvis within last 12 months	
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 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 tumor -primary respiratory diseases other than COPD (e.g., asthma) - impaired mobility related to non-COPD 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 Classification	- history of COPD ≥GOLD III - impaired mobility related to non-CHF causes, as judged by the investigator

Table 4.2: Summary of participant demographics and clinical characteristics. † Highlights acronyms used in table for: Montreal Cognitive Assessment (MoCA), Late-Life Functional and Disability Instrument (LLFDI), Unified Parkinson's Disease Rating Scale (UPDRS), Hoehn and Yahr Score (H&Y), Expanded Disability Status Scale (EDSS), Short Physical Performance Battery (SPPB), COPD Assessment Test (CAT), Forced expiratory volume (FEV1), Forced vital capacity (FVC) six-minute walk test (6MWT) and Kansas City Cardiomyopathy Questionnaire (KCCQ)

		Characterisation of groups					
		HA (n = 20)	**PD (n = 20)**	**MS (n = 20)**	**PFF (n = 19)**	**COPD (n = 17)**	**CHF (n = 12)**
Generic characteristics	**Sex (Male/Female)**	11/9	16/4	11/9	8/11	9/8	8/4
	Age (years)	71.7±5.8	69.8±7.2	48.7±9.7	80.0±8.5	69.4±9.1	69.1±11.7
	Height (m)	1.66±0.10	1.73±0.07	1.71±0.13	1.69±0.08	1.69±0.07	1.74±0.10
	Body Mass (Kg)	75.1±11.8	78.2±14.4	84.0±22.9	68.4±16.0	73.7±14.2	84.5±16.8
Cognition	**MoCA†**	27.7±2.6	24.6±4.0	26.7±3.1	24.1±4.2	24.6±3.4	27.1±2.9
No. of Fallers	**Had a fall in last 12 months**	3	10	11	19	1	3
Pain - VAS Score (0, no pain – 100 worst)	**General**	11.1±18.6	20.5±26.4	23.4±24.7	13.2±16.9	14.1±16.1	16.8±28.1
	When Walking	8.1±13.9	21.8±27.2	26.4±32.5	25.5±27.4	13±14.4	17.8±30.1
No. of Walking Aid Users	**General Use**	1	6	5	13	1	4
	Laboratory Protocol	0	1	3	6	0	4
Cohort-specific Outcomes	**LLFDI†**	73.53±14.22	60.26±12.51	57.34±10.66	52.59±16.61	59.07±7.96	67.29±21.35
	UPDRS III†	-	28.4±13.6	-	-	-	-
	H&Y Score†	-	I n=4, II n=11, III n=5	-	-	-	-
	EDSS†	-	-	3.5±1.7	-	-	-
	SPPB†	-	-	-	6.2±3.9	-	-
	CAT score† (0, best – 40, worst)	-	-	-	-	16.6±8.9	-
	FEV_1† (L)	-	-	-	-	1.6±0.6	-
	FVC† (L)	-	-	-	-	2.9±0.7	-
	6MWT† Distance (m)	-	-	-	-	357.6±88.5	370.7±115.6
	KCCQ†	-	-	-	-	-	80.5±20.2

4.2.2 Experimental Protocol

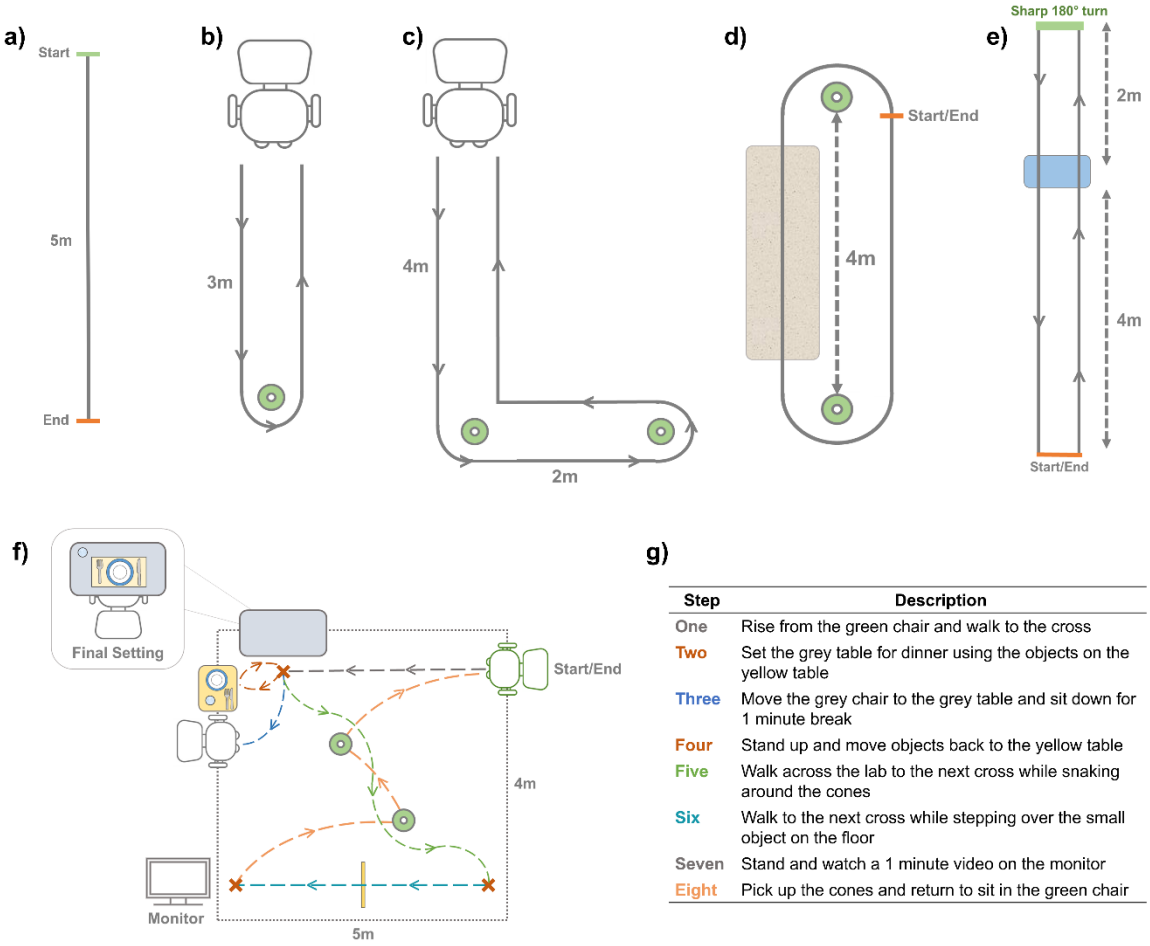
Laboratory-based Assessment

Based on previous literature (Ainsworth et al., 1993; Bourke et al., 2017), five key elements associated with walking in real-world scenarios were identified as necessary to vary in the multi-task protocol: speed, incline/steps, surface, path shape, and cognitive demand. In addition, specific motor tasks and postures that may abruptly alter the participants strategy of walking were included to further broaden the simulation of typical real-world transitions (e.g., walk-to-sit).

Besides including a large variation in walking paths and transitions, a critical target for the analysis was that the desired DMOs could be calculated for each WB. The focus was specifically placed on WS as a summary measure of these walking variations. In the context of this study, a WB was defined as a period of walking that included at least two consecutive strides for each leg (Kluge et al., 2021). When considering normative values for stride length at approximately 1.1-1.5 m (Hollman et al., 2011), this definition safely translates into a minimal travelled distance of 3.5 m for each period of walking. For the laboratory-based assessment, given the limitations of the 3D motion capture systems within the five gait laboratories involved in the study, the protocol capture area was designed to be smaller than 5 m×4 m.

In light of the above considerations, the designed protocol included seven structured tasks with each task varying in at least one of the identified elements of a walking path that are subject to change in real-world scenarios, as well as variation in postural transitions at the start and end of each task. In addition, a task that focused on simulating daily activities was designed, that resulted in the most complex combinations of walking paths and transitions (Mazzà et al., 2021). The eight tasks were performed, within in the capture volume of the motion capture systems, in the order presented below. As variation in walking was pertinent to the success of this task protocol, it was decided that tasks would only be performed once for each participant, except for the straight walking trials that were performed twice due to its high importance in the initial stages of the validation. All tasks were described and demonstrated to the participant by the researcher prior to completing the task, allowing the participant the opportunity to determine if they felt comfortable and safe performing the task. In addition, adjustments (e.g., whether the participant wished to use their walking aid and/or modulation to the task difficulty, as described below) was determined by the participant and researcher at this stage:

- *Straight Walking*: Participants were asked to walk a predefined path of 5 m from a standing start to a standing end. This trial was performed twice at three self-selected speeds: comfortable, slow, and fast (Figure 4.1a).
- *Timed Up and Go (TUG)*: At a comfortable speed, participants were asked to rise from the chair and walk 3 m to the cone, make a 180° left hand turn around the cone, walk back to chair and sit down (Figure 4.1b). The TUG is a standard clinical assessment of mobility for patients with varying level of mobility. Although the task setup conforms to the published guidelines (Podsiadlo & Richardson, 1991a), the primary reason for including this task is the sit-to-walk/walk-to-sit posture and turn.
- *L-Test*: At a comfortable speed, participants were asked to rise from the chair and walk 4 m to the first cone, make a 90° turn to the left around the cone and walk straight to the second cone, turn 180° to the left around the cone and walk straight before making a final 90° turn to the right of the first cone and walking back to sit in the chair (Figure 4.1c). Like the TUG, this test has previously been validated as a clinical assessment in patient cohorts (Deathe & Miller, 2005) however, the main purpose of including this test in the proposed protocol is the variation of turns included in a short walkway.
- *Surface Test*: Participants were asked to walk at a comfortable speed around the circuit by turning around the cones and stepping over the carpeted mat completing the circuit twice. Participants started and stopped the tasks in a standing position (Figure 4.1d).
- *Hallway Test*: From a standing start, participants were asked to walk at a comfortable speed along the predefined walkway stepping up and down from the step. At the end of the walkway participants completed a sharp 180° turn and walked back along the walkway (again stepping up and down the step) before coming to a stop at the end of the walkway (Figure 4.1e).
- *Simulated Daily Activities*: Participants were asked to start sitting in the green chair (Figure 4.1f) and complete a series of tasks defined in Figure 4.1g. while moving around the room. The tasks were split into separate steps, with the next set of instructions only given to the participant after the previous step had been completed. All steps for this task were completed at the preferred walking speed of the participant.



Step	Description
One	Rise from the green chair and walk to the cross
Two	Set the grey table for dinner using the objects on the yellow table
Three	Move the grey chair to the grey table and sit down for 1 minute break
Four	Stand up and move objects back to the yellow table
Five	Walk across the lab to the next cross while snaking around the cones
Six	Walk to the next cross while stepping over the small object on the floor
Seven	Stand and watch a 1 minute video on the monitor
Eight	Pick up the cones and return to sit in the green chair

Figure 4.1. Schematics of the seven structured tasks: a) straight walking, b) TUG, c) L-test, d) Surface test and e) Hallway test completed in the laboratory-based assessment, as well as a schematic of the simulated daily activities f) and description of the eight steps performed during this task g).

This multi-task design allowed for a tiered approach to the data collection, with tasks set in ascending order of complexity (based on consensus of the authors) to allow participants to ask to stop the data collection at any point that the protocol became too burdensome. The grade of difficulty within each task was also modulated to accommodate the use of walking aids, arm rests for the TUG and L-test, handrails for the step in the Hallway test and by removing obstacles from the Simulated Daily Activities path to facilitate clearance. This tiered approach was deemed appropriate to account for the wide variation in physical health and level of mobility across the sample populations to ensure that a meaningful amount of data could be collected for all participants while safeguarding the participants' safety and well-being.

During all the above tasks, a 3D motion capture system was used to record the trajectories of two markers located on each foot (Heel and Toe) and a four-marker cluster on the lower back (Figure 4.2). The temporal and spatial parameters needed to compute WS were calculated using the marker trajectories for each detected WB, as described in Bonci et al., (2022).

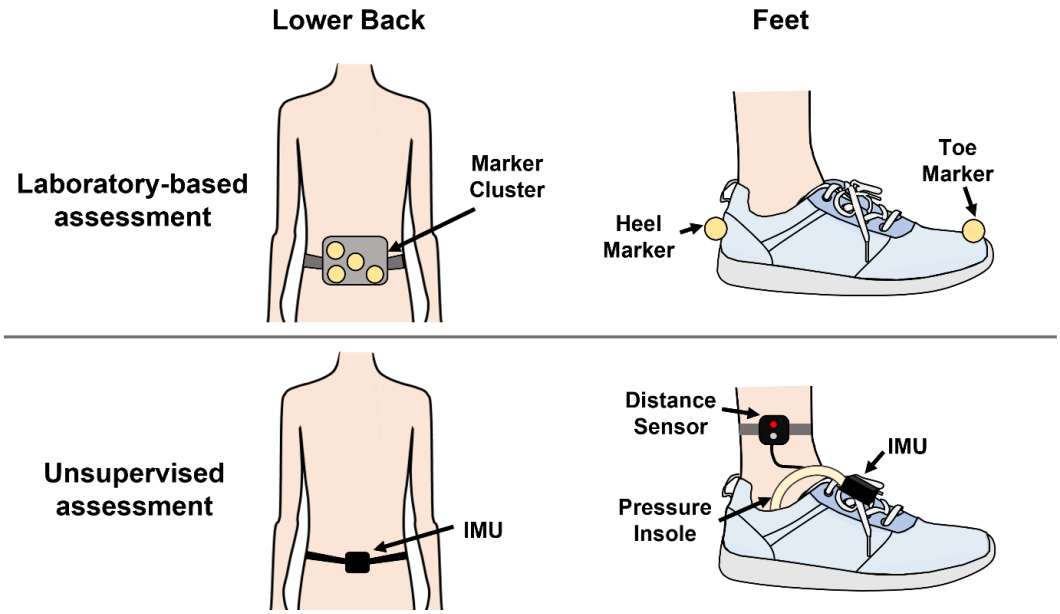


Figure 4.2. Illustration of the participant setup in the laboratory assessment and the multi-sensor system worn during the unsupervised assessment, consisting of IMUs on the lower back and feet, as well as 16-point pressure insole in each shoe and time-of-flight distance sensor on both legs (at a comfortable height above the medial malleolus).

Unsupervised Assessment

In addition to the laboratory-based assessment, all participants underwent unsupervised data collection in their habitual environment. During this session participants were asked to go about their typical routine and consider incorporating some specific tasks to ensure the presence of some variability/crucial elements in the data collection: outdoor walking; walking an inclined or declined path; moving from one room to another; walking up and down stairs. WS was recorded using a multi-sensor wearable system (INertial module with DIstance sensors and Pressure insoles, INDIP) integrating multiple IMUs, 16-points pressure insoles and infrared time-of-flight distance sensors (Figure 4.2) (Salis, Bertuletti, Bonci, et al., 2021; Salis, Bertuletti, Scott, et al., 2021). Gait events were detected by the pressure insoles and foot IMUs of the INDIP system and combined, with priority given to the events detected by the pressure insoles if detected by both methods (Salis, Bertuletti, Bonci, et al., 2021), which allowed the calculation of the temporal parameters for each detected WB.

Spatial parameters were then calculated from the foot IMUs data based on the direct and reverse integration approach described in (Salis, Bertuletti, Scott, et al., 2021). The duration of this session was set to 2.5 hours to minimise participant burden and safely operate within the limits imposed by the battery of the INDIP system.

When compared to the marker trajectory method used in the laboratory-based assessment, the INDIP system estimated walking speed with a median absolute error of 0.02m/s (Salis, Bertuletti, Bonci, et al., 2021; Salis, Bertuletti, Scott, et al., 2021)*. It was hence deemed appropriate to use bins of 0.1m/s for comparing the data obtained from the different systems used in the laboratory and in the real-world observation.*

4.2.3 Data Analysis

Once WBs were identified for each task and participant, the relevant WS was calculated as the average stride speed over the considered WB.

$$\boldsymbol{WS} = \frac{\sum_{\boldsymbol{k=1}}^{\boldsymbol{n_{strides}}} \boldsymbol{StrideSpeed_k}}{\boldsymbol{n_{strides}}}$$

Where $n_{strides}$is the number of strides identified in a WB and $StrideSpeed_k$ is the walking speed of stride k in the WB, defined as:

$$\boldsymbol{StrideSpeed_k}\ [\boldsymbol{^m/_s}] = \frac{\boldsymbol{StrideLength_k}\ [\boldsymbol{m}]}{\boldsymbol{StrideDuration_k}\ [\boldsymbol{s}]}$$

When more than one WB was detected within a task (as expected with the Simulated Daily Activities), the WS for all WBs belonging to that task was included. The frequency and distribution of the WS recorded in the lab were then computed for each cohort.

To establish whether the 2.5 hours of unsupervised recording was sufficient to reach an adequate sample size from a statistical point of view, a preliminary statistical power analysis was conducted to define the minimum number of WBs for each cohort required to validate the estimate of WS. In order to obtain a confidence interval smaller than 0.1 when comparing two different instruments with an α=0.05 and a power ß=0.9, an analysis was performed in Stata 16.1 (Stata Corp LP; College Station, Texas, USA; command line: sampicc 0.7 2, alpha (0.05) power (0.9) w (0.1) ci), which showed that the minimum number of WBs needed in each cohort for an Intraclass Correlation Coefficient ICC≥0.7 would be 401.

Finally, to determine whether the laboratory-based protocol could mimic the same WS range as recorded in the 2.5-hour task, the minimum and maximum walking speeds were extracted for all WBs recorded both in the supervised and unsupervised testing. The bias and limits of agreement of these two variables were then calculated and used to create Bland Altman plots for each cohort.

4.3 Results

4.3.1 Protocol Safety and Feasibility

The protocol proved to be safe and feasible. No adverse events were recorded, despite the fact that it was administered to patients with severe mobility impairments and during the COVID-19 pandemic. From the 108 participants included, 100% managed to complete the straight walking at a comfortable speed, the TUG, and the surface test; >95% completed the slow and fast walking tasks and the L-test, and>85% also completed the Hallway test and the Simulated Daily Activities (see details in

Figure 4.3). Participants that were not able to complete some of tasks were always from a patient cohort and were generally reported to be more severely affected by their individual disease based on the cohort specific outcomes.

The design of the protocol allowed calculation of WS and other DMOs in the vast majority of the recorded tasks, with the exceptions being mostly in the fast speed walking (

Figure 4.3), where the limited available space did not allow the recording of enough strides to satisfy the WB definition criteria. Those participants for which DMOs were not calculated were primarily within the healthy older adult population or participants who exhibited milder disease severity as indicated by the lower clinical scores.

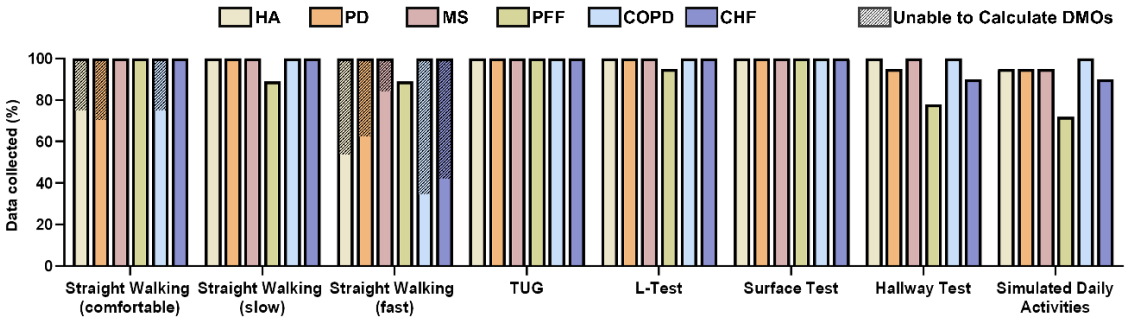


Figure 4.3. Data successfully collected for each laboratory-based task in percentage of planned. Striped portions of the bar are from trials in which data was collected but not enough strides were collected to identify a WB and calculate WS.

As per the third protocol design objective of inducing a variety of gait patterns, a large range of WSs were captured for each cohort (Figure 4.4). In particular, the addition of the more complex tasks allowed a much greater spread in WSs to be achieved compared to the standard straight walking trials. The distribution of the WS recorded for the PFF patients was skewed by the lowest speeds, but overall, the protocol allowed the same range of speeds to be observed for the other groups.

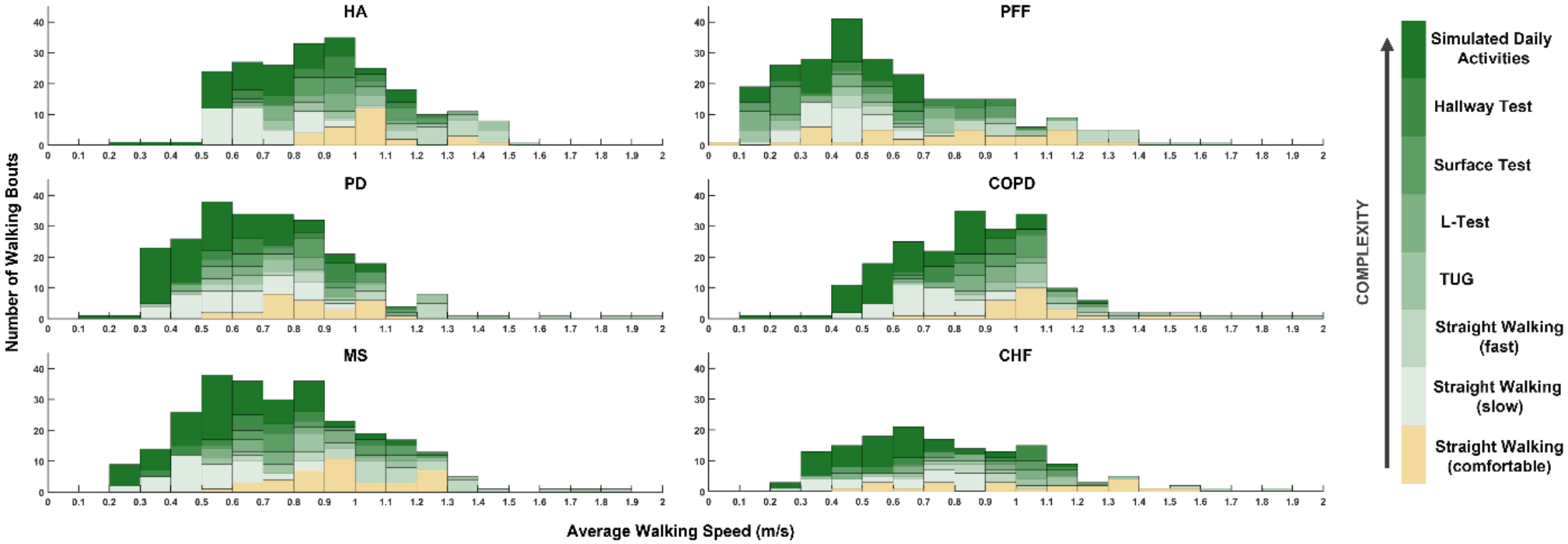


Figure 4.4 Stacked bars representing the distribution and frequency of the average walking speeds collected for the laboratory-based protocol. The colouring represents the contribution from individual tasks to the overall number of WB recorded at each speed, with orange bars highlighting the Straight Walking tasks at a comfortable speed. The green bars are the other tasks, with the darker colour related to the higher complexity of the task. Data used to generate the graphs is available as Supplementary material

4.3.2 Duration of Real-World Observation

From the 108 participants recruited in this data collection, data from 102 were included in the analysis of the real-world observation (HA=20, PD=18, MS=19, PFF=16, COPD=17 and CHF=12). The reduced number was the result of technical complications with the synchronisation of the multiple systems used or experimental error during the observation. Figure 4.5 shows the total number of walking bouts recorded for each participant and cohort. In total, these were 1330 for the HA, 678 for the PD, 771 for the MS, 628 for the PFF, 1035 for the COPD and 696 for the CHF cohort.

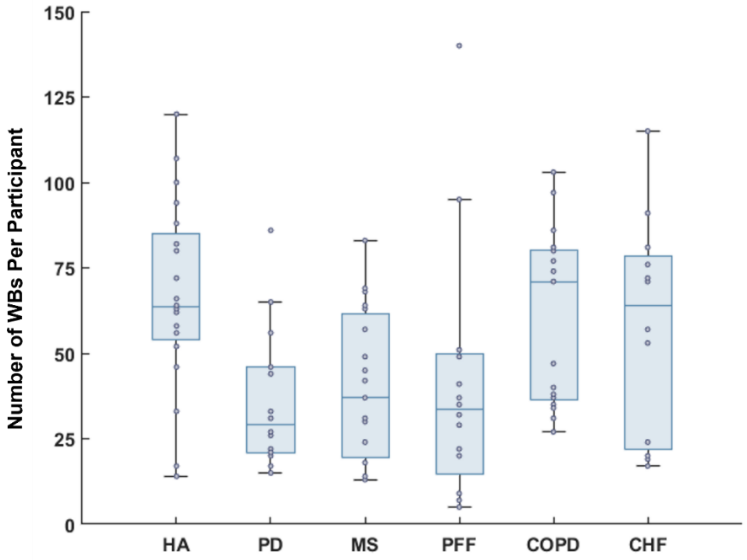


Figure 4.5 Box plot of the total number of WBs recorded for each cohort, with data points for each participant, during the 2.5-hour unsupervised data collection in the participants' habitual environment.

At participant level, the range of speeds observed in the lab was smaller than the one observed in the unsupervised assessment in the real-world (Table 4.3). Figure 4.6 shows the LOA and bias of the minimum (blue) and maximum (red) walking speeds collected in the laboratory versus the values recorded during the 2.5 hours for each participant. The minimum and maximum walking speeds observed in the real-world were generally slower than those recorded in the laboratory. However, these differences varied between cohorts, with the largest mean bias observed across all groups being 0.3m/s for the maximum walking speed of PFF. In all groups, the limits of agreement for the maximum speeds were bigger than those observed for the minimum speeds.

Table 4.3: Descriptive statistics of the walking speed captured in the laboratory-based assessment and 2.5 hour of unsupervised monitoring in the participants habitual environment for each cohort.

Walking Speed (m/s)										
Cohort	**Laboratory-based assessment**					**2.5-hour unsupervised monitoring**				
	Mean	**STD**	**25-perc**	**Median**	**75-perc**	**Mean**	**STD**	**25-perc**	**Median**	**75-perc**
OHA (n=20)	0.83	0.30	0.61	0.85	1.04	0.63	0.29	0.42	0.56	0.79
PD (n=19)	0.70	0.30	0.50	0.66	0.86	0.66	0.34	0.39	0.61	0.85
MS (n=19)	0.77	0.30	0.54	0.73	0.98	0.70	0.32	0.46	0.64	0.87
PFF (n=16)	0.60	0.33	0.37	0.52	0.80	0.50	0.21	0.35	0.48	0.63
COPD (n=17)	0.80	0.31	0.59	0.82	1.01	0.61	0.22	0.46	0.59	0.74
CHF (n=11)	0.78	0.33	0.52	0.71	1.00	0.79	0.34	0.53	0.77	1.06

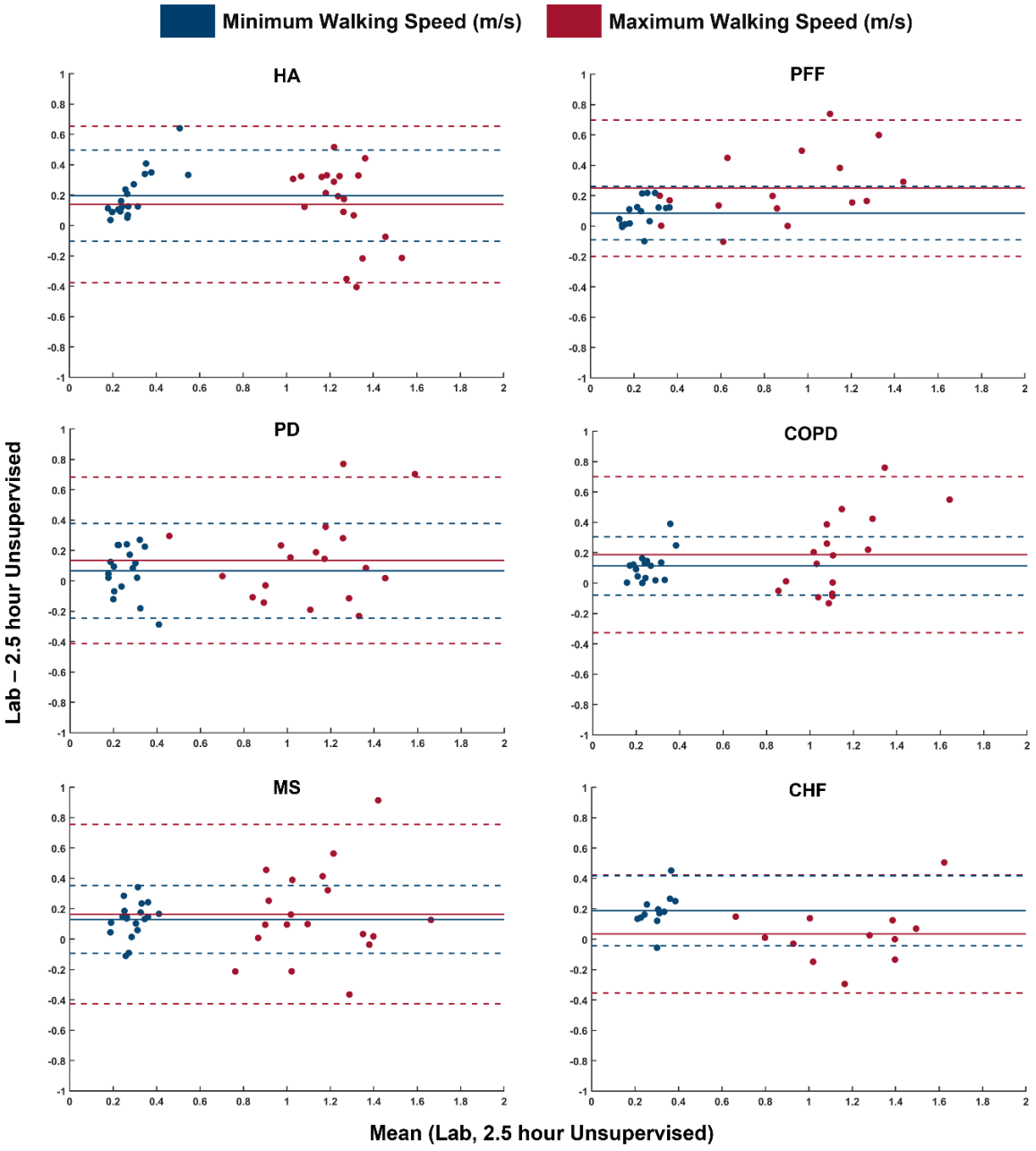


Figure 4.6 Bland-Altman plots of the minimum (blue) and maximum (red) walking speed recorded in the laboratory-based protocol compared to 2.5 hours of unsupervised monitoring in a habitual environment. The solid horizontal lines (-) indicate the mean bias and dashed horizontal lines (--) the upper and lower limits of agreement (LOA). Data used to generate the graphs is available as Supplementary material.

4.4 Discussion

This paper validated a protocol designed for simulating real-world gait in a laboratory setting. The reported results show that the proposed protocol was successfully designed and met its objectives. All participants were able to complete the majority (five out of seven) of the laboratory-based tasks, with the data collection typically lasting about 45 minutes. The selected tasks allowed for a broad range of WB speeds to be recorded in all groups, irrespective of their gait impairments. Furthermore, the 2.5 hours unsupervised observation in the participants' habitual environment allowed recording a number of WBs that was higher than the 401 threshold needed to ensure reliability and validity of a device for real-world gait monitoring in a given cohort.

A novelty introduced in the protocol was the tiered approach to increase difficulty with each new task while allowing for some adaptations to task setup to ensure inclusivity and safety of all participants (e.g., by allowing use of walking aids or armrests), with the Hallway test and Simulated Daily Activities being the most challenging. This ensured that the more severely disabled participants could opt out of tasks if they felt tired, unsafe, or too challenged, while also allowing for a variety of data to be collected. This proved to be a necessary and effective approach: the participants unable to complete some of the more complex tasks were all within the groups from a patient group and were rather severely affected by their individual disease based on the cohort specific outcomes. Notably, previous studies adopting complex protocols similarly succeeded in collecting a variety of gait data only included a few groups of participants, primarily older healthy adults (Allet et al., 2008; Lee & Park, 2011; O'Brien et al., 2019) and individuals with PD (Allet et al., 2008; O'Brien et al., 2019). The successful inclusion of participants presenting with different types of pathological gait, is a clear indication of the protocol's suitability to be safely and effectively deployed also in vulnerable groups.

The results showed that increasing task complexity allowed broad ranges of WS to be captured for all participants, in comparison to only including tasks at natural speed, as is commonly the case in validation studies (Aich et al., 2018; Han et al., 2019b; Mico-Amigo et al., 2016). This was particularly evident for the HA and COPD cohort, where most participants walked at very similar speeds (HA 0.8 – 1.2 m/s and COPD 0.9 – 1.2 m/s) in the straight walking tests. Interestingly, asking the participants to perform the most complex tasks, namely the Hallway test and the Simulated Daily Activities, proved more effective at inducing slower walking speeds than simply asking participants to walk slowly. This can be explained by the increase in motor and cognitive demand associated with the tasks, in which the majority of walking periods included a change in walking environment (i.e., obstacles to avoid or stepover on the ground) or some form of dual task (i.e., carrying an

object and walking), which are known to induce a speed reduction both in healthy and pathological groups (Smith et al., 2015) in more complex or noisy environments (Kowalsky et al., 2021).

When leaving the laboratory and moving to the unsupervised observations in the habitual environment, the open question was around the effectiveness of the protocol in providing enough WBs to ensure a robust validation of the WS estimates. Setting the duration to 2.5-hours and requesting that participants to consider incorporating some specific tasks within their unsupervised activities proved to be effective in all groups as reflected by the number of WBs recorded at group level. This was despite the different group sizes and the range of mobility disability exhibited. In general, as expected, the participants with the higher levels of severity in regard to their condition were those for whom the lower number of WBs were recorded.

When comparing data from the laboratory and the real-world sessions, the minimum and maximum WS results highlighted an overall bias with higher WSs recorded in the laboratory. Although this may not be surprising, the mean bias for both maximum and minimum walking speeds for all, but the maximum walking speed for PFF, fell below 0.2m/s. Though this bias can be considered low compared to the larger biases reported in the literature for healthy adults, Parkinson's disease, and multiple sclerosis (Shah et al., 2020; Warmerdam et al., 2020) it is likely explained by the fact that previous studies have included a much longer period of observation (e.g., 7 days data). While this aspect does not affect the usefulness of the 2.5-hours observation for the purposes of gathering enough validation data, it is certainly an indicator of the likelihood that longer periods of continuous monitoring are needed when interested in assessing an individual's mobility performance. Further studies would be required to further investigate this aspect and establish the link between capacity and performance assessments.

The results come with their limitations. The unbalanced size of the observed groups, which resulted from recruiting extremely vulnerable groups during the COVID-19 pandemic did not hinder our ability to achieve the study's primary aims, but it certainly prevented further investigation on the differences between the groups and on the effect of specific disease severity. From a more practical perspective, another limitation sits in the fact that in the context of this multicentric study the walkway length (5m) was dictated by the 3D motion capture system volume. The confinements to a small space did not allow the recording of more than four consecutive strides in the same direction: this is even more evident during the straight walks, in particular for participants with higher walking speeds and longer stride length, rather than during complex tasks including curvilinear portions, which cover

longer distances in the capture volume. We would certainly recommend future studies consider extending this walkway length.

4.5 Conclusion

This study presents an innovative multi-task gait assessment protocol beyond straight walking. It allows a relatively realistic representation of daily life relevant mobility aspects and can therefore be used for the validation of monitoring devices used in real life. Of particular note is the suitability of the protocol for measuring gait in conditions typically associated with pathological gait. This suggests that it can also be used to detect changes in gait due to, for example, the onset or progression of a disease, or due to therapy. Ultimately, this protocol opens up the option of capturing entirely new aspects of gait in real life, such as balance control in response to obstacles, directional behaviour, and cognitive aspects of mobility

5 DISCUSSION

A part of Mobilise-D, a large European project, the work carried out within this PhD contributed to developing and implementing a digital mobility assessment solution to demonstrate that digital mobility outcomes (DMOs) can successfully predict clinical outcomes and assist regulatory and clinical decision making.

Mobilise-D aims to validate this digital mobility assessment in different patient populations where measures of mobility loss can be related to disease progression. The validation process is being achieved via two separate stages: a technical validation study, aiming at establishing accuracy and reliability of selected DMOs, and a clinical validation study aiming at proving their construct validity, predictive capacity, and ability to detect change. As part of the technical validation study, the specific aim of this thesis was to define, validate and deploy an experimental protocol and the relevant algorithmic and laboratory tools needed to enable a robust multicentric collection of gait data from a wide range of disease groups.

The above aims were to be achieved through the completion of the three following main objectives:

1. Assess and review previous protocols that aim to test the technical readiness of a wearable sensor as a real-world digital mobility assessment, via a scoping review (Chapter 2).
2. Identify and develop the software and tools needed to ensure reproducibility of experimental procedures in a multicentric study and verify the reliability and robustness of the above tools (Chapter 3).
3. Define and validate a laboratory based experimental protocol to validate the algorithms used for calculating DMOs. This was achieved by evaluating the effectiveness of the protocol and looking into its ability to induce a large variety of walking speeds, effectively mimicking real world gait (Chapter 4).

While the main results and limitations of the individual studies have already been discussed in the relevant chapters, the focus here is to highlight the link between them, the impact of the individual findings and the future investigations that have already stemmed from this thesis.

The completion of a scoping review on the current approaches to assess the technical readiness of a digital mobility assessment for real-world gait highlighted that although there was an abundance of previous efforts, the heterogeneity of the protocols, sample

groups and outcomes validated, meant that there was no clear validation framework, subsequently limiting the comparability of results between studies. Nonetheless, the state of the art at the time of the review highlighted concepts that were crucial in the development of a validation. In particular, it clearly emerged that a robust and thorough simulation of real-world conditions would imply the need for a multi context and multi-stage protocol and that this should be possible to administer to an array of participant groups. Since completing this review, the literature has since been updated to include the years the between 2020 and 2022. It was found that there had been limited changes in validation protocols and less articles in general. This was most likely linked to the restrictions to experimental studies deriving from the Covid-19 pandemic during this timeframe. In response to the restriction in experimental work during this time, there has been a focus on methodological papers that are using an ethos to frame future validations in a similar manner to that proposed in this thesis. This includes the RADAR-AD study that aims to validate remote monitoring of gait in patients with Alzheimer's, which has incorporated a multi-stage and multi-context task protocol (Muurling et al., 2021).

Although this review was extensive in its search, and aimed to be inclusive of all research, it is important to mention some limitations in this work. First, the screening of the papers was completed only by myself, which in the practice of literature review is suggested as to be avoided to reduce bias from the article screening (Page et al., 2021). Regardless, I believe that the articles that were found gave a generally good understanding of the current state of the art with a clear saturation in concepts and themes found in this search that indicate that all relevant sources were found and assessed. This is further supported by strong similarities in findings to a recent review published (Peters et al., 2021) that considered the same topic to the one presented in this thesis.

Based on the scoping review, it was evident that although a lot of research has been conducted in this field, there is large level of heterogeneity between studies, meaning that collation and comparison of results is not easy. As such, what is required when wanting to assess the technical readiness of a wearable sensor to estimate real-world walking, is the development of a framework of standardised tools and experimental procedures to enable direct and safe comparison of data.

Following this literature review, it emerged clearly that an optoelectronic system would be the ideal tool to be used as a gold standard for the evaluation of DMOs obtained using wearable devices. Nonetheless, as amply discussed in Chapter 3, a validation based on the comparison to a gold standard must entail the standardisation of the procedures for collecting and processing the reference data. Although optoelectronic stereophotogrammetry is a regularly used reference system when evaluating inertial

measurement units, data coming from these systems are not free of inaccuracies. Accounting for the concurrent possible sources of error, it is crucial that a system is reviewed in its configuration of use and experimental spot-checks and data pre-processing tools are put in place to minimise the propagation of error. The framework and relevant software and tools developed in this thesis and described in chapter 3 proved to be successful when deployed within the mobilise-D project. Notably, within the project significant attention has been put into ensuring that every step of the data collection and analysis would satisfy the ALCOA+ standards (EMA, 2010): not only all data should be data be attributable, legible, contemporaneous, original, and accurate but also complete, consistent, enduring, and available. This is a standard recognised by the pharmacological companies and by the regulatory bodies, which the tools here proposed allowed to fully implement. Indeed, the design of the technical validation and of the procedures here adopted as part of a letter of support received from the European Medical Agency, EMA (Viceconti et al., 2022).

Last but not least, a clear impact for the broader research community was achieved with the publication and sharing of the results through conferences and papers and by making the software needed to implement the suggested procedures openly available via Figshare, with 497 views and 272 downloads at the time of writing this.

A general limitation of this work was the restrictions to their implementation within one study only. However, with the software publicly available there is opportunity for this work to be tailored to other studies to maintain the same standard for the reference system. In addition, with the continued push from multicentric work which is crucial in this area these tools can easily be used across sites, across systems and across users.

The results from Chapter 2 and Chapter 3 were combined into a multi-stage task protocol, that was successfully deployed in five different centres and to five different diseased groups. As detailed in Chapter 4, the task protocol was well received with no adverse events noted throughout the study.

Besides achieving all of its initial objectives, the richness of the data that the protocol allowed to collect is expected to enable the exploration of other factors such as variations between the motor strategies adopted by the different cohorts in executing the various tasks. For example, the characterisation of the turning strategies adopted by the different cohorts is now being further investigated. Gait in real life is complex – tripping, falling, turning – challenging for patients because disease specific alternation in gait might make it difficult. These concepts are particularly relevant for the MS, PD, and PFF cohorts, sharing the general theme of unsteady gait that subsequently leads to falls. This common tendency to increased risk of falls might however be due to different disease-specific impairment:

ataxia is a common clinical symptom in patients (80%) with Multiple sclerosis and predominantly presents as tremors or loss of coordination and unsteadiness in gait. In Parkinson's disease shuffling and, with disease severity, festination and freezing of gait are common symptoms that could lead to a fall. Finally, in PFF this could be explained by asymmetry and reduced range of motion. A protocol entailing different types of turning like the one here proposed is expected to enable further analysis of these aspects. This could indeed be investigated by looking into how walking speed is modulated across the different motor tasks considering both spatial and temporal parameters and identifying cohort specific gait strategies in coping with transitions that challenges balance. A concrete example of how this could be investigated is shown in Figure 5.1 below, where the strategies adopted by the healthy adult and cohort of participants with multiple sclerosis have been summarised in terms of modulation of stride length duration and width (represented by inter-heel distance) to achieve the different observed walking speed, as recorded during straight walking (SW), turns of approximately 90° (CW1) and turns of approximately 180° (CW2). From this preliminary analysis, it appears there is a higher variation in stride length and stride width in the patient cohort when compared to healthy adult controls, particularly in the wider turns.

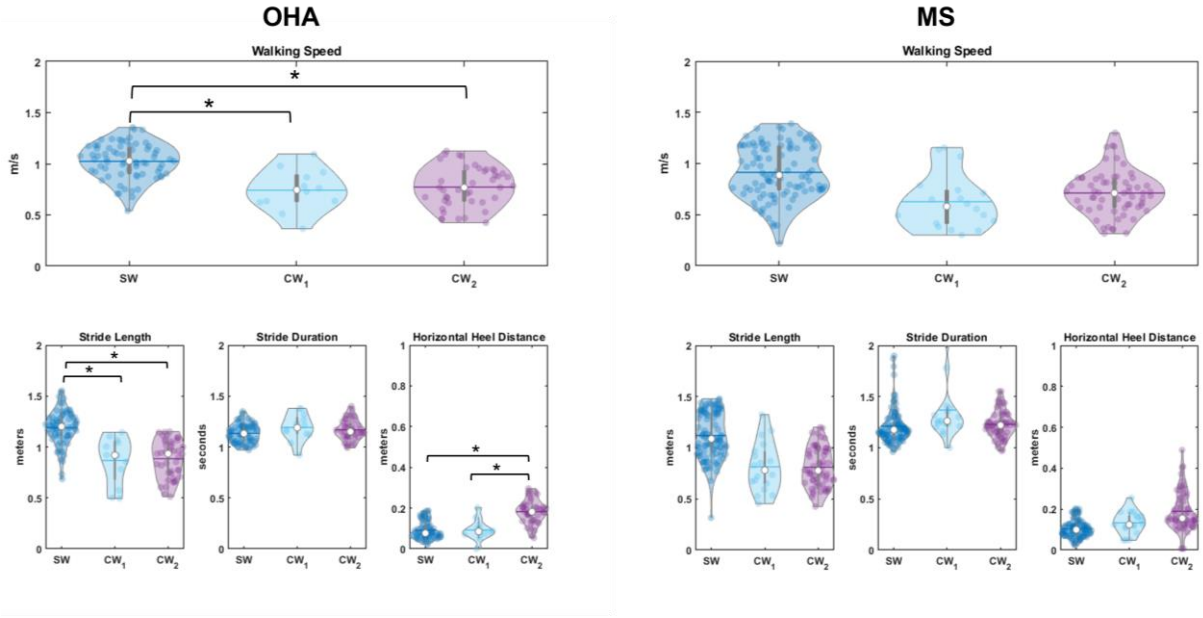


Figure 5.1 Violin plots of the walking speed, stride length, stride duration and horizontal heel distance seen for stragith walking (SW), turns of approximately 90° (CW1) and turns of approximately 180° (CW2) for older healthy adults (OHA) and multiple sclerosis cohort (MS).

A further aspect that could significantly contribute to the understanding of the risk of falls is the assessment of the toe clearance. This has been further investigated as part of a side project enabled by this thesis, the details of which are described in (please see Appendix

Four). This project entailed the development of a new methodology to improve standard methods for the calculation of toe clearance based on the digitisation of the shoe perimeter. The initial findings from this study allowed to identify differences between cohorts when applied to the data collected with the protocol.

5.1 Conclusion

In conclusion, from this thesis, it feels clear the objectives defined for the PhD have been met successfully, and not only have they meant success for the project but their acceptance, outreach and use by the research community hopefully shows the initial need and success of this work. The future direction of this work has not only meant the acceptance form the research community, but it has allowed the progression of the Mobilise-D project to the clinical validation, which includes 2,500 patients from 30+ clinical sites. The work from this thesis and general developments of the technical validation helped to shape the framework, training and resources for this study and the algorithms validated at the technical validations will be implemented on larger scale in the clinical validation.

REFERENCES

Abdel-Aziz, Y. I., & Karara, H. M. (2015). Direct Linear Transformation from Comparator Coordinates into Object Space Coordinates in Close-Range Photogrammetry. *Photogrammetric Engineering & Remote Sensing*, *81*(2), 103-107. https://doi.org/10.14358/pers.81.2.103

Aich, S., Pradhan, P. M., Park, J., Sethi, N., Vathsa, V. S. S., & Kim, H.-C. (2018). A Validation Study of Freezing of Gait (FoG) Detection and Machine-Learning-Based FoG Prediction Using Estimated Gait Characteristics with a Wearable Accelerometer. *Sensors (Basel)*, *18*(10), 3287. https://doi.org/10.3390/s18103287

Ainsworth, B. E., Haskell, W. L., Leon, A. S., Jacobs, D. R., Montoye, H. J., Sallis, J. F., & Paffenbarger Jr, R. S. (1993). Compendium of physical activities: Classification of energy costs of human physical activities. *Med Sci Sports Exerc*, *25*(1), 71-80. https://doi.org/10.1249/00005768-199301000-00011

Allet, L., Armand, S., de Bie, R. A., Golay, A., Monnin, D., Aminian, K., & de Bruin, E. D. (2008). Reliability of diabetic patients' gait parameters in a challenging environment. *Gait Posture*, *28*(4), 680-686. https://doi.org/10.1016/j.gaitpost.2008.05.006

Aminian, K., Najafi, B., Büla, C., Leyvraz, P. F., & Robert, P. (2002). Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes. *J Biomech*, *35*(5), 689-699. https://doi.org/10.1016/S0021-9290(02)00008-8

Aminian, K., Rezakhanlou, K., De Andres, E., Fritsch, C., Leyvraz, P. F., & Robert, P. (1999). Temporal feature estimation during walking using miniature accelerometers: An analysis of gait improvement after hip arthroplasty. *Med Biol Eng Comput*, *37*(6), 686-691. https://doi.org/10.1007/BF02513368

APDM. (2021). *Comprehensive gait and balance analysis*. https://apdm.com/mobility/

Atallah, L., Wiik, A., Lo, B., Cobb, J. P., Amis, A. A., & Yang, G. Z. (2014). Gait asymmetry detection in older adults using a light ear-worn sensor. *Physiol. Meas*, *35*(5), N29-N40. https://doi.org/10.1088/0967-3334/35/5/N29

Aurand, A. M., Dufour, J. S., & Marras, W. S. (2017). Accuracy map of an optical motion capture system with 42 or 21 cameras in a large measurement volume. *Journal of biomechanics.*, *58*, 237-240. https://doi.org/10.1016/j.jbiomech.2017.05.006

Baker, R., Leboeuf, F., Reay, J., & Sangeux, M. (2018). The Conventional Gait Model - Success and Limitations. In *Handbook of Human Motion* (pp. 489-508). Springer International Publishing. https://doi.org/10.1007/978-3-319-14418-4_25

Barnett, A., & Cerin, E. (2006). Individual calibration for estimating free-living walking speed using the MTI monitor. *Med Sci Sports Exerc*, *38*(4), 761-767. https://doi.org/10.1249/01.mss.0000210206.55941.b2

Barre, A., & Armand, S. (2014). Biomechanical ToolKit: Open-source framework to visualize and process biomechanical data. *Comput Methods Programs Biomed*, *114*(1), 80-87. https://doi.org/10.1016/j.cmpb.2014.01.012

Bejarano, N. C., Ambrosini, E., Pedrocchi, A., Ferrigno, G., Monticone, M., & Ferrante, S. (2014). A novel adaptive, real-time algorithm to detect gait events from wearable sensors. *IEEE transactions on neural systems and rehabilitation engineering*, *23*(3), 413-422.

Ben Mansour, K., Rezzoug, N., & Gorce, P. (2015). Foot side detection from lower lumbar spine acceleration. *Gait Posture*, *42*(3), 386-389. https://doi.org/10.1016/j.gaitpost.2015.05.021

Bertuletti, S., Della Croce, U., & Cereatti, A. (2019). A wearable solution for accurate step detection based on the direct measurement of the inter-foot distance. *J Biomech*, *84*, 274-277. https://doi.org/10.1016/j.jbiomech.2018.12.039

Bobilev, A., Bolotin, Y. V., Voronov, A., & Kruchinin, P. (2012). About two the least-squares method modifications for lost data recovery in videoanalysis system based on accelerometer data. *Russ J Biomech*, *16*(1), 80-91.

Bonci, T., Salis, F., Scott, K., Alcock, L., Becker, C., Bertuletti, S., . . . Mazza, C. (2022). An Algorithm for Accurate Marker-Based Gait Event Detection in Healthy and Pathological Populations During Complex Motor Tasks. *Frontiers in bioengineering and biotechnology*, *10*, 868928-868928. https://doi.org/10.3389/fbioe.2022.868928

Booth, M. (2000). Assessment of Physical Activity: An International Perspective. *Research Quarterly for Exercise and Sport*, *71*(sup2), 114-120. https://doi.org/10.1080/02701367.2000.11082794

Bourke, A. K., Ihlen, E. A. F., Bergquist, R., Wik, P. B., Vereijken, B., & Helbostad, J. L. (2017). A physical activity reference data-set recorded from older adults using body-worn inertial sensors and video technology—The ADAPT study data-set. *Sensors (Basel)*, *17*(3), 559. https://doi.org/10.3390/s17030559

Brodie, M. A. D., Coppens, M. J. M., Lord, S. R., Lovell, N. H., Gschwind, Y. J., Redmond, S. J., . . . Delbaere, K. (2016). Wearable pendant device monitoring using new wavelet-based methods shows daily life and laboratory gaits are different. *Med Biol Eng Comput*, *54*(4), 663-674. https://doi.org/10.1007/s11517-015-1357-9

Buckley, C., Alcock, L., McArdle, R., Ur Rehman, R. Z., Del Din, S., Mazzà, C., . . . Rochester, L. (2019). The role of movement analysis in diagnosing and monitoring neurodegenerative conditions: Insights from gait and postural control. *Brain Sci*, *9*(2), 34. https://doi.org/10.3390/brainsci9020034

Byun, S., Han, J. W., Kim, T. H., & Kim, K. W. (2016). Test-retest reliability and concurrent validity of a single tri-axial accelerometer-based gait analysis in older adults with normal cognition. *PLoS One*, *11*(7), e0158956-e0158956. https://doi.org/10.1371/journal.pone.0158956

Camargo, J., Ramanathan, A., Csomay-Shanklin, N., & Young, A. (2020). Automated gap-filling for marker-based biomechanical motion capture data. *Computer Methods in Biomechanics and Biomedical Engineering*, *23*(15), 1180-1189. https://doi.org/10.1080/10255842.2020.1789971

Cappello, A., Cappozzo, A., La Palombara, P. F., Lucchetti, L., & Leardini, A. (1997). Multiple anatomical landmark calibration for optimal bone pose estimation. *Human movement science*, *16*(2-3), 259-274. https://doi.org/10.1016/S0167-9457(96)00055-3

Cappozzo, A., Della Croce, U., Leardini, A., & Chiari, L. (2005). Human movement analysis using stereophotogrammetry. *Gait & posture.*, *21*(2), 186-196. https://doi.org/10.1016/j.gaitpost.2004.01.010

Chan, C. W., & Rudins, A. (1994). Foot biomechanics during walking and running. Mayo Clinic Proceedings,

Chiari, L., Croce, U. D., Leardini, A., & Cappozzo, A. (2005). Human movement analysis using stereophotogrammetry: Part 2: Instrumental errors. *Gait Posture*, *21*(2), 197-211. https://doi.org/10.1016/j.gaitpost.2004.04.004

Cho, Y.-S., Jang, S.-H., Cho, J.-S., Kim, M.-J., Lee, H. D., Lee, S. Y., & Moon, S.-B. (2018). Evaluation of validity and reliability of inertial measurement unit-based gait analysis systems. *Ann Rehabil Med*, *42*(6), 872-883. https://doi.org/10.5535/arm.2018.42.6.872

Conconi, M., Pompili, A., Sancisi, N., & Parenti-Castelli, V. (2021). Quantification of the errors associated with marker occlusion in stereophotogrammetric systems and implications on gait analysis. *J Biomech*, *114*, 110162-110162. https://doi.org/10.1016/j.jbiomech.2020.110162

Costa, T. M., Simieli, L., Bersotti, F. M., Mochizuki, L., Barbieri, F. A., & Coelho, D. B. (2022). Gait and posture are correlated domains in Parkinson's disease. *Neuroscience letters.*, *775*, 136537. https://doi.org/10.1016/j.neulet.2022.136537

Crapo, R. O., Casaburi, R., Coates, A. L., Enright, P. L., MacIntyre, N. R., McKay, R. T., . . . Mottram, C. (2002). ATS statement: Guidelines for the six-minute walk test. *Am J Respir Crit Care Med*, *166*(1), 111-117. https://doi.org/10.1164/ajrccm.166.1.at1102

Davis, R. B., Õunpuu, S., Tyburski, D., & Gage, J. R. (1991). A gait analysis data collection and reduction technique. *Human movement science*, *10*(5), 575-587. https://doi.org/10.1016/0167-9457(91)90046-Z

De Vroey, H., Staes, F., Weygers, I., Vereecke, E., Vanrenterghem, J., Deklerck, J., . . . Claeys, K. (2018). The implementation of inertial sensors for the assessment of temporal parameters of gait in the knee arthroplasty population. *Clin Biomech (Bristol, Avon)*, *54*, 22-27. https://doi.org/10.1016/j.clinbiomech.2018.03.002

Deathe, A. B., & Miller, W. C. (2005). 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*, *85*(7), 626-635. https://doi.org/10.1093/ptj/85.7.626

Del Din, S., Godfrey, A., Galna, B., Lord, S., & Rochester, L. (2016). Free-living gait characteristics in ageing and Parkinson's disease: Impact of environment and ambulatory bout length. *J Neuroeng Rehabil*, *13*(1), 46-46. https://doi.org/10.1186/s12984-016-0154-5

Del Din, S., Godfrey, A., Mazza, C., Lord, S., & Rochester, L. (2016). Free-living monitoring of Parkinson's disease: lessons from the field.

Del Din, S., Godfrey, A., & Rochester, L. (2016). Validation of an Accelerometer to Quantify a Comprehensive Battery of Gait Characteristics in Healthy Older Adults and Parkinson's Disease: Toward Clinical and at Home Use. *IEEE J Biomed Health Inform*, *20*(3), 838-847. https://doi.org/10.1109/JBHI.2015.2419317

Del Din, S., Hickey, A., Woodman, S., Hiden, H., Morris, R., Watson, P., . . . Godfrey, A. (2016, 2016-06-01). Accelerometer-based gait assessment: Pragmatic deployment on an international scale. 2016 IEEE Statistical Signal Processing Workshop (SSP),

Della Croce, U., & Cappozzo, A. (2000). A spot check for estimating stereophotogrammetric errors. *Med Biol Eng Comput*, *38*(3), 260-266. https://doi.org/10.1007/bf02347045

DeLuzio, K. J., Wyss, U. P., Li, J., & Costigan, P. A. (1993). A procedure to validate three-dimensional motion assessment systems. *J Biomech*, *26*(6), 753-759. https://doi.org/10.1016/0021-9290(93)90037-F

Derungs, A., Schuster-Amft, C., & Amft, O. (2018). Longitudinal walking analysis in hemiparetic patients using wearable motion sensors: Is there convergence between body sides? *Front Bioeng Biotechnol*, *6*(MAY), 57-57. https://doi.org/10.3389/fbioe.2018.00057

Di Marco, R., Rossi, S., Castelli, E., Patanè, F., Mazzà, C., & Cappa, P. (2017). Effects of the calibration procedure on the metrological performances of stereophotogrammetric systems for human movement analysis. *Measurement : journal of the International Measurement Confederation*, *101*, 265-271. https://doi.org/10.1016/j.measurement.2016.01.008

Diaz Novo, C., Alharbi, S., Fox, M., Ouellette, E., Biden, E., Tingley, M., & Chester, V. (2014). The impact of technical parameters such as video sensor technology, system configuration, marker size and speed on the accuracy of motion analysis systems. *Ingeniería mecánica, tecnología y desarrollo*, *5*(1), 265-271.

Dijkstra, B., Zijlstra, W., Scherder, E., & Kamsma, Y. (2008). Detection of walking periods and number of steps in older adults and patients with Parkinson's disease: accuracy of a pedometer and an accelerometry-based method. *Age Ageing*, *37*(4), 436-441. https://doi.org/10.1093/ageing/afn097

Djuric-Jovicic, M., Jovicic, N., Radovanovic, S., Kresojevic, N., Kostic, V., & Popovic, M. (2014). Quantitative and qualitative gait assessments in Parkinson's disease patients. *Vojnosanit Pregl*, *71*(9), 809-816. https://doi.org/10.2298/VSP130422012D

Dobkin, B. H., Xu, X., Batalin, M., Thomas, S., & Kaiser, W. (2011). Reliability and validity of bilateral ankle accelerometer algorithms for activity recognition and walking speed after stroke. *Stroke*, *42*(8), 2246-2250. https://doi.org/10.1161/STROKEAHA.110.611095

Ehara, Y., Fujimoto, H., Miyazaki, S., Mochimaru, M., Tanaka, S., & Yamamoto, S. (1997). Comparison of the performance of 3D camera systems II. *Gait & posture*, *5*(3), 251-255. https://doi.org/10.1016/S0966-6362(96)01093-4

Eichelberger, P., Ferraro, M., Minder, U., Denton, T., Blasimann, A., Krause, F., & Baur, H. (2016). Analysis of accuracy in optical motion capture–A protocol for laboratory setup evaluation. *J Biomech*, *49*(10), 2085-2088. https://doi.org/10.1016/j.jbiomech.2016.05.007

EMA. (2010). *Reflection paper on expectations for electronic source data and data transcribed to electronic data collection tools in clinical trials*.

Espay, A. J., Hausdorff, J. M., Sanchez-Ferro, A., Klucken, J., Merola, A., Bonato, P., . . . Maetzler, W. (2019). A roadmap for implementation of patient-centered digital outcome measures in Parkinson's disease obtained using mobile health technologies. *Mov Disord*, *34*(5), 657-663. https://doi.org/10.1002/mds.27671

Everaert, D. G., Spaepen, A. J., Wouters, M. J., Stappaerts, K. H., & Oostendorp, R. A. B. (1999). Measuring small linear displacements with a three-dimensional video motion analysis system: Determining its accuracy and precision. *Arch Phys Med Rehabil*, *80*(9), 1082-1089. https://doi.org/10.1016/S0003-9993(99)90065-5

Fasel, B., Duc, C., Dadashi, F., Bardyn, F., Savary, M., Farine, P.-A., & Aminian, K. (2017). A wrist sensor and algorithm to determine instantaneous walking cadence and speed in daily life walking. *Med Biol Eng Comput*, *55*(10), 1773-1785. https://doi.org/10.1007/s11517-017-1621-2

Ferrari, A., Ginis, P., Hardegger, M., Casamassima, F., Rocchi, L., & Chiari, L. (2015). A mobile Kalman-filter based solution for the real-time estimation of spatio-temporal gait parameters. *IEEE transactions on neural systems and rehabilitation engineering*, *24*(7), 764-773.

Figueiredo, J., Felix, P., Costa, L., Moreno, J. C., & Santos, C. P. (2018). Gait event detection in controlled and real-life situations: Repeated measures from healthy subjects. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(10), 1945-1956.

Fourati, H. (2015). Heterogeneous Data Fusion Algorithm for Pedestrian Navigation via Foot-Mounted Inertial Measurement Unit and Complementary Filter. *IEEE transactions on instrumentation and measurement*, *64*(1), 221-229. https://doi.org/10.1109/TIM.2014.2335912

Genovese, V., Mannini, A., & Sabatini, A. M. (2017). A Smartwatch Step Counter for Slow and Intermittent Ambulation. *IEEE access*, *5*, 13028-13037. https://doi.org/10.1109/ACCESS.2017.2702066

Gill, T. M., Allore, H. G., Hardy, S. E., & Guo, Z. (2006). The Dynamic Nature of Mobility Disability in Older Persons. *J Am Geriatr Soc*, *54*(2), 248-254. https://doi.org/10.1111/j.1532-5415.2005.00586.x

Givon, U., Zeilig, G., & Achiron, A. (2009). Gait analysis in multiple sclerosis: Characterization of temporal–spatial parameters using GAITRite functional ambulation system. *Gait & posture.*, *29*(1), 138-142. https://doi.org/10.1016/j.gaitpost.2008.07.011

Godfrey, A., Del Din, S., Barry, G., Mathers, J. C., & Rochester, L. (2015). Instrumenting gait with an accelerometer: A system and algorithm examination. *Med Eng Phys*, *37*(4), 400-407. https://doi.org/10.1016/j.medengphy.2015.02.003

Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., . . . LaPelle, N. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov. Disord*, *23*(15), 2129-2170. https://doi.org/10.1002/mds.22340

González, R. C., López, A. M., Rodriguez-Uría, J., Álvarez, D., & Alvarez, J. C. (2010). Real-time gait event detection for normal subjects from lower trunk accelerations. *Gait & posture.*, *31*(3), 322-325. https://doi.org/10.1016/j.gaitpost.2009.11.014

Han, J., Jeon, H. S., Yi, W. J., Jeon, B. S., & Park, K. S. (2009). Adaptive windowing for gait phase discrimination in Parkinsonian gait using 3-axis acceleration signals. *Med Biol Eng Comput*, *47*(11), 1155-1164. https://doi.org/10.1007/s11517-009-0521-5

Han, Y. C., Wong, K. I., & Murray, I. (2019a). Gait Phase Detection for Normal and Abnormal Gaits Using IMU. *IEEE sensors journal*, *19*(9), 3439-3448. https://doi.org/10.1109/JSEN.2019.2894143

Han, Y. C., Wong, K. I., & Murray, I. (2019b). Stride Length Estimation Based on a Single Shank's Gyroscope. *IEEE sensors letters*, *3*(10), 1-4. https://doi.org/10.1109/LSENS.2019.2942177

Hannink, J., Kautz, T., Pasluosta, C. F., Gasmann, K.-G., Klucken, J., & Eskofier, B. M. (2017). Sensor-Based Gait Parameter Extraction With Deep Convolutional Neural Networks. *IEEE J Biomed Health Inform*, *21*(1), 85-93. https://doi.org/10.1109/JBHI.2016.2636456

Hao, M., Chen, K., & Fu, C. (2019). Smoother-Based 3-D Foot Trajectory Estimation Using Inertial Sensors. *IEEE Trans Biomed Eng*, *66*(12), 3534-3542. https://doi.org/10.1109/TBME.2019.2907322

Hegde, N., Bries, M., Swibas, T., Melanson, E., & Sazonov, E. (2018). Automatic Recognition of Activities of Daily Living Utilizing Insole-Based and Wrist-Worn Wearable Sensors. *IEEE J Biomed Health Inform*, *22*(4), 979-988. https://doi.org/10.1109/JBHI.2017.2734803

Hickey, A., Del Din, S., Rochester, L., & Godfrey, A. (2016). Detecting free-living steps and walking bouts: validating an algorithm for macro gait analysis. *Physiol. Meas*, *38*(1), N1-N15. https://doi.org/10.1088/1361-6579/38/1/N1

Hollman, J. H., McDade, E. M., & Petersen, R. C. (2011). Normative spatiotemporal gait parameters in older adults. *Gait Posture*, *34*(1), 111-118. https://doi.org/10.1016/j.gaitpost.2011.03.024

Holtzer, R., Verghese, J., Xue, X., & Lipton, R. B. (2006). Cognitive Processes Related to Gait Velocity: Results From the Einstein Aging Study. *Neuropsychology*, *20*(2), 215-223. https://doi.org/10.1037/0894-4105.20.2.215

Howarth, S. J., & Callaghan, J. P. (2010). Quantitative assessment of the accuracy for three interpolation techniques in kinematic analysis of human movement. *Computer Methods in Biomechanics and Biomedical Engineering*, *13*(6), 847-855. https://doi.org/10.1080/10255841003664701

Hsing-Cheng, C., Yu-Liang, H., Shih-Chin, Y., Jung-Chih, L., & Zhi-Hao, W. (2016). A Wearable Inertial Measurement System With Complementary Filter for Gait Analysis of Patients With Stroke or Parkinson's Disease. *IEEE access*, *4*, 8442-8453. https://doi.org/10.1109/ACCESS.2016.2633304

Hundza, S. R., Hook, W. R., Harris, C. R., Mahajan, S. V., Leslie, P. A., Spani, C. A., . . . Livingston, N. J. (2013). Accurate and reliable gait cycle detection in Parkinson's disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *22*(1), 127-137.

Hwang, T.-H., Reh, J., Effenberg, A. O., & Blume, H. (2018). Real-Time Gait Analysis Using a Single Head-Worn Inertial Measurement Unit. *IEEE transactions on consumer electronics*, *64*(2), 240-248. https://doi.org/10.1109/TCE.2018.2843289

Jarchi, D., Wong, C., Kwasnicki, R. M., Heller, B., Tew, G. A., & Yang, G.-Z. (2014). Gait Parameter Estimation From a Miniaturized Ear-Worn Sensor Using Singular Spectrum Analysis and Longest Common Subsequence. *IEEE Trans Biomed Eng*, *61*(4), 1261-1273. https://doi.org/10.1109/TBME.2014.2299772

Jette, A. M., Haley, S. M., Coster, W. J., Kooyoomjian, J. T., Levenson, S., Heeren, T., & Ashba, J. (2002). Late Life Function and Disability Instrument: I. Development and Evaluation of the Disability Component. *The journals of gerontology.*, *57*(4), M209-M216. https://doi.org/10.1093/gerona/57.4.m209

Ji, Q., Yang, L., Li, W., Zhou, C., & Ye, X. (2018). Real-time gait event detection in a real-world environment using a laser-ranging sensor and gyroscope fusion method. *Physiol. Meas*, *39*(12), 125003-125003. https://doi.org/10.1088/1361-6579/aae7ee

Jones, D. S., & Podolsky, S. H. (2015). The history and fate of the gold standard. *Lancet*, *385*(9977), 1502-1503. https://doi.org/10.1016/S0140-6736(15)60742-5

Khalaf, K., Hulleck, A. A., Menoth Mohan, D., Abdallah, N., & El-Rich, M. (2022). Present and future of gait assessment in clinical practice: Towards the application of novel trends and technologies. *Frontiers in Medical Technology*, 91.

Kluge, F., Del Din, S., Cereatti, A., Gaßner, H., Hansen, C., Helbostad, J. L., . . . Mazzà, C. (2021). Consensus based framework for digital mobility monitoring. *PLoS One*, *16*(8), e0256541-e0256541. https://doi.org/10.1371/journal.pone.0256541

Kowalsky, D. B., Rebula, J. R., Ojeda, L. V., Adamczyk, P. G., & Kuo, A. D. (2021). Human walking in the real world: Interactions between terrain type, gait parameters, and energy expenditure. *PLoS One*, *16*(1), e0228682-e0228682. https://doi.org/10.1371/journal.pone.0228682

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology*, *33*(11), 1444-1444. https://doi.org/10.1212/wnl.33.11.1444

Kuxhaus, L., Schimoler, P. J., Vipperman, J. S., & Miller, M. C. (2009). Effects of Camera Switching on Fine Accuracy in a Motion Capture System. *J Biomech Eng*, *131*(1), 014502-014506. https://doi.org/10.1115/1.3002910

Küderle, A., Roth, N., Zlatanovic, J., Zrenner, M., Eskofier, B., & Kluge, F. (2022). The placement of foot-mounted IMU sensors does affect the accuracy of spatial parameters during regular walking. *PLOS ONE*, *17*(6), e0269567. https://doi.org/10.1371/journal.pone.0269567

Lee, J.-A., Cho, S.-H., Lee, Y.-J., Yang, H.-K., & Lee, J.-W. (2010). Portable Activity Monitoring System for Temporal Parameters of Gait Cycles. *J Med Syst*, *34*(5), 959-966. https://doi.org/10.1007/s10916-009-9311-8

Lee, J. K., & Park, E. J. (2011). Quasi real-time gait event detection using shank-attached gyroscopes. *Med Biol Eng Comput*, *49*(6), 707-712. https://doi.org/10.1007/s11517-011-0736-0

Lee, M., Youm, C., Jeon, J., Cheon, S.-M., & Park, H. (2018). Validity of shoe-type inertial measurement units for Parkinson's disease patients during treadmill walking. *J Neuroeng Rehabil*, *15*(1), 38-38. https://doi.org/10.1186/s12984-018-0384-9

Liu, G., & Mcmillan, L. (2006). Estimation of missing markers in human motion capture. *The Visual Computer*, *22*(9-11), 721-728. https://doi.org/10.1007/s00371-006-0080-9

Lord, S., Galna, B., Verghese, J., Coleman, S., Burn, D., & Rochester, L. (2013). Independent domains of gait in older adults and associated motor and nonmotor attributes: Validation of a factor analysis approach. *J Gerontol A Biol Sci Med Sci*, *68*(7), 820-827. https://doi.org/10.1093/gerona/gls255

Lord, S., Galna, B., Verghese, J., Coleman, S., Burn, D., & Rochester, L. (2013). Independent Domains of Gait in Older Adults and Associated Motor and Nonmotor Attributes: Validation of a Factor Analysis Approach. *The journals of gerontology.*, *68*(7), 820-827. https://doi.org/10.1093/gerona/gls255

Lowe, S. A., & ÓLaighin, G. (2014). Monitoring human health behaviour in one's living environment: A technological review. *Medical engineering & physics.*, *36*(2), 147-168. https://doi.org/10.1016/j.medengphy.2013.11.010

Maetzler, W., Klucken, J., & Horne, M. (2016). A clinical view on the development of technology-based tools in managing Parkinson's disease. *Movement Disorders*, *31*(9), 1263-1271. https://doi.org/10.1002/mds.26673

Maletsky, L. P., Sun, J., & Morton, N. A. (2007). Short communication: Accuracy of an optical active-marker system to track the relative motion of rigid bodies. *Journal of biomechanics*, *40*(3), 682-685. https://doi.org/10.1016/j.jbiomech.2006.01.017

Maqbool, H. F., Husman, M. A. B., Awad, M. I., Abouhossein, A., & Dehghani-Sanij, A. A. (2015). Real-time gait event detection for transfemoral amputees during ramp ascending and descending. *EMBC*, *2015-*, 4785-4788. https://doi.org/10.1109/EMBC.2015.7319464

Mariani, B., Hoskovec, C., Rochat, S., Büla, C., Penders, J., & Aminian, K. (2010). 3D gait assessment in young and elderly subjects using foot-worn inertial sensors. *J Biomech*, *43*(15), 2999-3006. https://doi.org/10.1016/j.jbiomech.2010.07.003

Mariani, B., Jiménez, M. C., Vingerhoets, F. J. G., & Aminian, K. (2013). On-Shoe Wearable Sensors for Gait and Turning Assessment of Patients With Parkinson's Disease. *IEEE Trans Biomed Eng*, *60*(1), 155-158. https://doi.org/10.1109/TBME.2012.2227317

Martin, C. L., Phillips, B. A., Kilpatrick, T. J., Butzkueven, H., Tubridy, N., McDonald, E., & Galea, M. P. (2006). Gait and balance impairment in early multiple sclerosis in the absence of clinical disability. *Multiple sclerosis.*, *12*(5), 620-628. https://doi.org/10.1177/1352458506070658

Martindale, C. F., Roth, N., Gassner, H., List, J., Regensburger, M., Eskofier, B. M., & Kohl, Z. (2020). Technical Validation of an Automated Mobile Gait Analysis System for Hereditary Spastic Paraplegia Patients. *IEEE J Biomed Health Inform*, *24*(5), 1490-1499. https://doi.org/10.1109/JBHI.2019.2937574

Martinez-Hernandez, U., & Dehghani-Sanij, A. A. (2018). Adaptive Bayesian inference system for recognition of walking activities and prediction of gait events using wearable sensors. *Neural Networks*, *102*, 107-119.

Mazzà, C., Alcock, L., Aminian, K., Becker, C., Bertuletti, S., Bonci, T., . . . Rochester, L. (2021). Technical validation of real-world monitoring of gait: a multicentric observational study. *BMJ Open*, *11*(12), e050785-e050785. https://doi.org/10.1136/bmjopen-2021-050785

McCamley, J., Donati, M., Grimpampi, E., & Mazzà, C. (2012). An enhanced estimate of initial contact and final contact instants of time using lower trunk inertial sensor data. *Gait & posture.*, *36*(2), 316-318. https://doi.org/10.1016/j.gaitpost.2012.02.019

McGinnis, R. S., Mahadevan, N., Moon, Y., Seagers, K., Sheth, N., Wright, J. A., . . . Patel, S. (2017). A machine learning approach for gait speed estimation using skin-mounted wearable sensors: From healthy controls to individuals with multiple sclerosis. *PLoS One*, *12*(6), e0178366-e0178366. https://doi.org/10.1371/journal.pone.0178366

Merriaux, P., Dupuis, Y., Boutteau, R., Vasseur, P., & Savatier, X. (2017). A study of vicon system positioning performance. *Sensors (Basel)*, *17*(7), 1591. https://doi.org/10.3390/s17071591

Mico-Amigo, M. E., Kingma, I., Ainsworth, E., Walgaard, S., Niessen, M., van Lummel, R. C., & van Dieen, J. H. (2016). A novel accelerometry-based algorithm for the detection of step durations over short episodes of gait in healthy elderly. *J Neuroeng Rehabil*, *13*(1), 38-38. https://doi.org/10.1186/s12984-016-0145-6

Miller, C., Mulavara, A., & Bloomberg, J. (2002). A quasi-static method for determining the characteristics of a motion capture camera system in a “split-volume” configuration. *Gait Posture*, *16*(3), 283-287. https://doi.org/10.1016/S0966-6362(02)00019-X

Misu, S., Asai, T., Ono, R., Sawa, R., Tsutsumimoto, K., Ando, H., & Doi, T. (2017). Development and validity of methods for the estimation of temporal gait

parameters from heel-attached inertial sensors in younger and older adults. *Gait Posture*, *57*, 295-298. https://doi.org/10.1016/j.gaitpost.2017.06.022

Miyazaki, S. (1997). Long-term unrestrained measurement of stride length and walking velocity utilizing a piezoelectric gyroscope. *IEEE Trans Biomed Eng*, *44*(8), 753-759. https://doi.org/10.1109/10.605434

Morris, S., Morris, M. E., & Iansek, R. (2001). Reliability of Measurements Obtained With the Timed "Up & Go" Test in People With Parkinson Disease. *Phys Ther*, *81*(2), 810-818. https://doi.org/10.1093/ptj/81.2.810

Motoi, K., Taniguchi, S., Baek, M., Wakugawa, M., Sonoda, T., Yuji, T., . . . Yamakoshi, K.-I. (2012). Development of a wearable gait monitoring system for evaluating efficacy of walking training in rehabilitation. *Sensors and materials*, *24*(6), 359-373. https://doi.org/10.18494/SAM.2012.792

Munn, Z., Peters, M. D. J., Stern, C., Tufanaru, C., Mcarthur, A., & Aromataris, E. (2018). Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Medical Research Methodology*, *18*(1). https://doi.org/10.1186/s12874-018-0611-x

Murray, M. P., Drought, A. B., & Kory, R. C. (1964). Walking Patterns of Normal Men. *J Bone Joint Surg Am*, *46*(2), 335-360. https://doi.org/10.2106/00004623-196446020-00009

Muurling, M., de Boer, C., Kozak, R., Religa, D., Koychev, I., Verheij, H., . . . Visser, P. J. (2021). Remote monitoring technologies in Alzheimer's disease: design of the RADAR-AD study. *Alzheimers Res Ther*, *13*(1), 89-89. https://doi.org/10.1186/s13195-021-00825-4

Niswander, W., & Kontson, K. (2021). Evaluating the Impact of IMU Sensor Location and Walking Task on Accuracy of Gait Event Detection Algorithms. *Sensors*, *21*(12), 3989. https://doi.org/10.3390/s21123989

O'Brien, M. K., Hidalgo-Araya, M. D., Mummidisetty, C. K., Vallery, H., Ghaffari, R., Rogers, J. A., . . . Jayaraman, A. (2019). Augmenting Clinical Outcome Measures of Gait and Balance with a Single Inertial Sensor in Age-Ranged Healthy Adults. *Sensors (Basel)*, *19*(20), 4537. https://doi.org/10.3390/s19204537

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., . . . McGuinness, L. A. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ : British medical journal.*, n71. https://doi.org/10.1136/bmj.n71

Palmerini, L., Reggi, L., Bonci, T., Del Din, S., Micó-Amigo, M. E., Salis, F., . . . Chiari, L. (2023). Mobility recorded by wearable devices and gold standards: the Mobilise-D procedure for data standardization. *Scientific Data*, *10*(1). https://doi.org/10.1038/s41597-023-01930-9

Paraschiv-Ionescu, A., Newman, C., Carcreff, L., Gerber, C. N., Armand, S., & Aminian, K. (2019). Locomotion and cadence detection using a single trunk-fixed accelerometer: Validity for children with cerebral palsy in daily life-like conditions. *J Neuroeng Rehabil*, *16*(1), 24-24. https://doi.org/10.1186/s12984-019-0494-z

Perera, S., Patel, K. V., Rosano, C., Rubin, S. M., Satterfield, S., Harris, T., . . . Studenski, S. A. (2015). Gait Speed Predicts Incident Disability: A Pooled Analysis. *J Gerontol A Biol Sci Med Sci*, *71*(1), 63-71. https://doi.org/10.1093/gerona/glv126

Peters, D. M., O'Brien, E. S., Kamrud, K. E., Roberts, S. M., Rooney, T. A., Thibodeau, K. P., . . . Mohapatra, S. (2021). Utilization of wearable technology to assess gait and mobility post-stroke: a systematic review. *Journal of NeuroEngineering and Rehabilitation*, *18*(1). https://doi.org/10.1186/s12984-021-00863-x

Pham, M. H., Elshehabi, M., Haertner, L., Din, S. D., Srulijes, K., Heger, T., . . . Maetzler, W. (2017). 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*, *8*(SEP), 457-457. https://doi.org/10.3389/fneur.2017.00457

Piazza, S. J., Erdemir, A., Okita, N., & Cavanagh, P. R. (2004). Assessment of the functional method of hip joint center location subject to reduced range of hip motion. *Journal of biomechanics.*, *37*(3), 349-356. https://doi.org/10.1016/s0021-9290(03)00288-4

Podsiadlo, D., & Richardson, S. (1991). The Timed Up & Go:a test of basic functional mobility for frail elderly persons. *J Am Geriatr Soc*, *39*(2), 142-148.

Prateek, G. V., Mazzoni, P., Earhart, G. M., & Nehorai, A. (2020). Gait Cycle Validation and Segmentation Using Inertial Sensors. *IEEE Trans Biomed Eng*, *67*(8), 2132-2144. https://doi.org/10.1109/TBME.2019.2955423

Punt, M., van Alphen, B., van de Port, I. G., van Dieen, J. H., Michael, K., Outermans, J. C., & Wittink, H. (2014). Clinimetric properties of a novel feedback device for assessing gait parameters in stroke survivors. *J Neuroeng Rehabil*, *11*(1), 30-30. https://doi.org/10.1186/1743-0003-11-30

Qiu, S., Wang, Z., Zhao, H., & Hu, H. (2016). Using Distributed Wearable Sensors to Measure and Evaluate Human Lower Limb Motions. *IEEE transactions on instrumentation and measurement*, *65*(4), 939-950. https://doi.org/10.1109/TIM.2015.2504078

Rampp, A., Barth, J., Schuelein, S., Gassmann, K.-G., Klucken, J., & Eskofier, B. M. (2015). Inertial Sensor-Based Stride Parameter Calculation From Gait Sequences in Geriatric Patients. *IEEE Trans Biomed Eng*, *62*(4), 1089-1097. https://doi.org/10.1109/TBME.2014.2368211

Richards, J. G. (1999). The measurement of human motion: A comparison of commercially available systems. *Human movement science*, *18*(5), 589-602. https://doi.org/10.1016/S0167-9457(99)00023-8

Robertson, D. G. E., Caldwell, G. E., Hamill, J., Kamen, G., & Whittlesey, S. N. (2014). Research methods in biomechanics.

Rochester, L., Galna, B., Lord, S., Mhiripiri, D., Eglon, G., & Chinnery, P. F. (2014). Gait impairment precedes clinical symptoms in spinocerebellar ataxia type 6: Gait Ataxia Precedes Clinical Symptoms in SCA6. *Movement disorders*, *29*(2), 252-255. https://doi.org/10.1002/mds.25706

Rochester, L., Mazzà, C., Mueller, A., Caulfield, B., McCarthy, M., Becker, C., . . . Roubenoff, R. (2020). A roadmap to inform development, validation and approval of digital mobility outcomes: the Mobilise-D approach.

Romijnders, R., Warmerdam, E., Hansen, C., Schmidt, G., & Maetzler, W. (2022). A Deep Learning Approach for Gait Event Detection from a Single Shank-Worn IMU: Validation in Healthy and Neurological Cohorts. *Sensors*, *22*(10), 3859. https://doi.org/10.3390/s22103859

Rucco, R., Agosti, V., Jacini, F., Sorrentino, P., Varriale, P., De Stefano, M., . . . Sorrentino, G. (2017). Spatio-temporal and kinematic gait analysis in patients with Frontotemporal dementia and Alzheimer's disease through 3D motion capture. *Gait & posture.*, *52*, 312-317. https://doi.org/10.1016/j.gaitpost.2016.12.021

Sabatini, A. M., Martelloni, C., Scapellato, S., & Cavallo, F. (2005). Assessment of walking features from foot inertial sensing. *IEEE Trans Biomed Eng*, *52*(3), 486-494. https://doi.org/10.1109/TBME.2004.840727

Salarian, A., Russmann, H., Vingerhoets, F. J. G., Dehollain, C., Blanc, Y., Burkhard, P. R., & Aminian, K. (2004). Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring. *IEEE Trans Biomed Eng*, *51*(8), 1434-1443. https://doi.org/10.1109/TBME.2004.827933

Salis, F., Bertuletti, S., Bonci, T., Della Croce, U., Mazzà, C., & Cereatti, A. (2021). A method for gait events detection based on low spatial resolution pressure insoles data. *Journal of biomechanics*, *127*, 110687-110687. https://doi.org/10.1016/j.jbiomech.2021.110687

Salis, F., Bertuletti, S., Scott, K., Caruso, M., Bonci, T., Buckley, E., . . . Cereatti, A. (2021). A wearable multi-sensor system for real world gait analysis. *IEEE ENG MED BIO*, 7020-7023. https://doi.org/10.1109/EMBC46164.2021.9630392

Schimpl, M., Lederer, C., & Daumer, M. (2011). Development and validation of a new method to measure walking speed in free-living environments using the actibelt® platform. *PLoS One*, *6*(8), e23080-e23080. https://doi.org/10.1371/journal.pone.0023080

Scott, K., Bonci, T., Alcock, L., Buckley, E., Hansen, C., Gazit, E., . . . Mobilise, D. C. (2021). A Quality Control Check to Ensure Comparability of Stereophotogrammetric Data between Sessions and Systems. *Sensors (Basel)*, *21*(24), 8223. https://doi.org/10.3390/s21248223

Sejdic, E., Lowry, K. A., Bellanca, J., Perera, S., Redfern, M. S., & Brach, J. S. (2016). Extraction of Stride Events From Gait Accelerometry During Treadmill Walking. *IEEE J Transl Eng Health Med*, *4*, 1-11. https://doi.org/10.1109/JTEHM.2015.2504961

Selles, R. W., Formanoy, M. A., Bussmann, J. B., Janssens, P. J., & Stam, H. J. (2005). Automated estimation of initial and terminal contact timing using accelerometers; development and validation in transtibial amputees and controls. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *13*(1), 81-88.

Shah, V. V., McNames, J., Mancini, M., Carlson-Kuhta, P., Spain, R. I., Nutt, J. G., . . . Horak, F. B. (2020). Laboratory versus daily life gait characteristics in patients with multiple

sclerosis, Parkinson's disease, and matched controls. *J Neuroeng Rehabil*, *17*(1), 159-159. https://doi.org/10.1186/s12984-020-00781-4

Skog, I., Handel, P., Nilsson, J.-O., & Rantakokko, J. (2010). Zero-Velocity Detection-An Algorithm Evaluation. *IEEE Trans Biomed Eng*, *57*(11), 2657-2666. https://doi.org/10.1109/TBME.2010.2060723

Slater, A. A., Hullfish, T. J., & Baxter, J. R. (2018). The impact of thigh and shank marker quantity on lower extremity kinematics using a constrained model. *BMC musculoskeletal disorders*, *19*(1), 399-399. https://doi.org/10.1186/s12891-018-2329-7

Smith, E., Cusack, T., & Blake, C. (2015). The effect of a dual task on gait speed in community dwelling older adults: A systematic review and meta-analysis. *Gait Posture*, *44*, 250-258. https://doi.org/10.1016/j.gaitpost.2015.12.017

Soltani, A., Dejnabadi, H., Savary, M., & Aminian, K. (2020). Real-World Gait Speed Estimation Using Wrist Sensor: A Personalized Approach. *IEEE J Biomed Health Inform*, *24*(3), 658-668. https://doi.org/10.1109/JBHI.2019.2914940

Sosnoff, J. J., Weikert, M., Dlugonski, D., Smith, D. C., & Motl, R. W. (2011). Quantifying gait impairment in multiple sclerosis using GAITRite™ technology. *Gait & posture.*, *34*(1), 145-147. https://doi.org/10.1016/j.gaitpost.2011.03.020

Sprager, S., & Juric, M. B. (2018). Robust stride segmentation of inertial signals based on local cyclicity estimation. *Sensors (Basel)*, *18*(4), 1091. https://doi.org/10.3390/s18041091

Stagni, R., Fantozzi, S., Cappello, A., & Leardini, A. (2005). Quantification of soft tissue artefact in motion analysis by combining 3D fluoroscopy and stereophotogrammetry: a study on two subjects. *Clinical biomechanics.*, *20*(3), 320-329. https://doi.org/10.1016/j.clinbiomech.2004.11.012

info:doi/10.1016/j.clinbiomech.2004.11.012

Storm, F. A., Buckley, C. J., & Mazzà, C. (2016). Gait event detection in laboratory and real life settings: Accuracy of ankle and waist sensor based methods. *Gait Posture*, *50*, 42-46. https://doi.org/10.1016/j.gaitpost.2016.08.012

Storm, F. A., Nair, K. P. S., Clarke, A. J., Van der Meulen, J. M., & Mazzà, C. (2018). Free-living and laboratory gait characteristics in patients with multiple sclerosis. *PLoS One*, *13*(5), e0196463-e0196463. https://doi.org/10.1371/journal.pone.0196463

Studenski, S., Perera, S., Patel, K., Rosano, C., Faulkner, K., Inzitari, M., . . . Guralnik, J. (2011). Gait Speed and Survival in Older Adults. *JAMA*, *305*(1), 50-58. https://doi.org/10.1001/jama.2010.1923

Supratak, A., Datta, G., Gafson, A. R., Nicholas, R., Guo, Y., & Matthews, P. M. (2018). Remote monitoring in the home validates clinical gait measures for multiple sclerosis. *Front Neurol*, *9*, 561-561. https://doi.org/10.3389/fneur.2018.00561

Systems, V. M. (2023). Model parameters and static/dynamic processing. In.

Teufl, W., Lorenz, M., Miezal, M., Taetz, B., Fröhlich, M., & Bleser, G. (2018). Towards inertial sensor based mobile gait analysis: Event-detection and spatio-temporal parameters. *Sensors (Basel)*, *19*(1), 38. https://doi.org/10.3390/s19010038

Tomioka, K., Iwamoto, J., Saeki, K., & Okamoto, N. (2011). Reliability and Validity of the International Physical Activity Questionnaire (IPAQ) in Elderly Adults: The Fujiwara-kyo Study. *Journal of Epidemiology*, *21*(6), 459-465. https://doi.org/10.2188/jea.je20110003

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., . . . Straus, S. E. (2018). PRISMA Extension for Scoping Reviews (PRISMA-ScR) : Checklist and Explanation.

van der Kruk, E., & Reijne, M. M. (2018). Accuracy of human motion capture systems for sport applications state-of-the-art review. *Eur J Sport Sci*, *18*(6), 806-819. https://doi.org/10.1080/17461391.2018.1463397

van Oeveren, B. T., de Ruiter, C. J., Beek, P. J., Rispens, S. M., & van Dieën, J. H. (2018). An adaptive, real-time cadence algorithm for unconstrained sensor placement. *Med Eng Phys*, *52*(February), 49-58. https://doi.org/10.1016/j.medengphy.2017.12.007

Versi, E. (1992). "Gold standard" is an appropriate term. *BMJ*, *305*(6846), 187-187. https://doi.org/10.1136/bmj.305.6846.187-b

Viceconti, M., Tome, M., Dartee, W., Knezevic, I., Hernandez Penna, S., Mazzà, C., . . . Maetzler, W. (2022). On the use of wearable sensors as mobility biomarkers in the marketing authorization of new drugs: A regulatory perspective. *Frontiers in medicine*, 2844.

Vicon Motion Systems, L. (2023). *Fill gaps in trial data*. https://docs.vicon.com/display/Nexus212/Fill+gaps+in+trial+data

Vincenzo, G., Andrea, M., & Angelo, M. S. (2017). A Smartwatch Step Counter for Slow and Intermittent Ambulation. *IEEE access*, *5*, 13028-13037. https://doi.org/10.1109/ACCESS.2017.2702066

Wang, L., Sun, Y., Li, Q., & Liu, T. (2018). Estimation of Step Length and Gait Asymmetry Using Wearable Inertial Sensors. *IEEE sensors journal*, *18*(9), 3844-3851. https://doi.org/10.1109/JSEN.2018.2815700

Wang, Z., Zhao, H., Qiu, S., & Gao, Q. (2015). Stance-Phase Detection for ZUPT-Aided Foot-Mounted Pedestrian Navigation System. *IEEE/ASME transactions on mechatronics*, *20*(6), 3170-3181. https://doi.org/10.1109/TMECH.2015.2430357

Warmerdam, E., Hausdorff, J. M., Atrsaei, A., Zhou, Y., Mirelman, A., Aminian, K., . . . Maetzler, W. (2020). Long-term unsupervised mobility assessment in movement disorders. *Lancet Neurol*, *19*(5), 462-470. https://doi.org/10.1016/S1474-4422(19)30397-7

Warmerdam, E., Romijnders, R., Geritz, J., Elshehabi, M., Maetzler, C., Otto, J. C., . . . Hansen, C. (2021). Proposed mobility assessments with simultaneous full-body inertial measurement units and optical motion capture in healthy adults and

neurological patients for future validation studies: Study protocol. *Sensors (Basel, Switzerland)*, *21*(17), 5833. https://doi.org/10.3390/s21175833

Windolf, M., Götzen, N., & Morlock, M. (2008). Systematic accuracy and precision analysis of video motion capturing systems—exemplified on the Vicon-460 system. *J Biomech*, *41*(12), 2776-2780. https://doi.org/10.1016/j.jbiomech.2008.06.024

World Health, O. (2001). International classification of functioning, disability and health : ICF.

Yang, C.-C., & Hsu, Y.-L. (2010). A Review of Accelerometry-Based Wearable Motion Detectors for Physical Activity Monitoring. *Sensors*, *10*(8), 7772-7788. https://doi.org/10.3390/s100807772

Yang, C.-C., Hsu, Y.-L., Shih, K.-S., & Lu, J.-M. (2011). Real-time gait cycle parameter recognition using a wearable accelerometry system. *Sensors (Basel)*, *11*(8), 7314-7326. https://doi.org/10.3390/s110807314

Yang, P.-F., Sanno, M., Brüggemann, G.-P., & Rittweger, J. (2012). Evaluation of the performance of a motion capture system for small displacement recording and a discussion for its application potential in bone deformation in vivo measurements. *Proc Inst Mech Eng H*, *226*(11), 838-847. https://doi.org/10.1177/0954411912452994

Yang, S., & Li, Q. (2012). IMU-based ambulatory walking speed estimation in constrained treadmill and overground walking. *Comput Methods Biomech Biomed Engin*, *15*(3), 313-322. https://doi.org/10.1080/10255842.2010.534465

Zhang, H., Guo, Y., & Zanotto, D. (2019). Accurate ambulatory gait analysis in walking and running using machine learning models. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *28*(1), 191-202.

Zhang, H., Zhang, J., Zhou, D., Wang, W., Li, J., Ran, F., & Ji, Y. (2015). Axis-Exchanged Compensation and Gait Parameters Analysis for High Accuracy Indoor Pedestrian Dead Reckoning. *Journal of sensors*, *2015*, 1-13. https://doi.org/10.1155/2015/915837

Zhou, H., Ji, N., Samuel, O. W., Cao, Y., Zhao, Z., Chen, S., & Li, G. (2016). Towards real-time detection of gait events on different terrains using time-frequency analysis and peak heuristics algorithm. *Sensors (Basel)*, *16*(10), 1634-1634. https://doi.org/10.3390/s16101634

Zijlstra, W. (2004). Assessment of spatio-temporal parameters during unconstrained walking. *European Journal of Applied Physiology*, *92*(1-2), 39-44. https://doi.org/10.1007/s00421-004-1041-5

Zijlstra, W., & Hof, A. L. (1997). Displacement of the pelvis during human walking: experimental data and model predictions. *Gait & posture.*, *6*(3), 249-262. https://doi.org/10.1016/S0966-6362(97)00021-0

Zijlstra, W., & Hof, A. L. (2003). Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture*, *18*(2), 1-10. https://doi.org/10.1016/S0966-6362(02)00190-X

LIST OF APPENDICES AND SUPPLEMENTARY MATERIAL

Appendix One: Full list of papers included in the literature review

- Aich, S., Pradhan, P. M., Park, J., Sethi, N., Vathsa, V. S. S., & Kim, H.-C. (2018). A Validation Study of Freezing of Gait (FoG) Detection and Machine-Learning-Based FoG Prediction Using Estimated Gait Characteristics with a Wearable Accelerometer. *Sensors (Basel)*, *18*(10), 3287. https://doi.org/10.3390/s18103287
- Allet, L., Armand, S., de Bie, R. A., Golay, A., Monnin, D., Aminian, K., & de Bruin, E. D. (2008). Reliability of diabetic patients' gait parameters in a challenging environment. *Gait Posture*, *28*(4), 680-686. https://doi.org/10.1016/j.gaitpost.2008.05.006
- Allseits, E., Agrawal, V., Lučarević, J., Gailey, R., Gaunaurd, I., & Bennett, C. (2018). A practical step length algorithm using lower limb angular velocities. *J Biomech*, *66*, 137-144. https://doi.org/10.1016/j.jbiomech.2017.11.010
- Allseits, E., Lučarević, J., Gailey, R., Agrawal, V., Gaunaurd, I., & Bennett, C. (2017). The Development and Concurrent Validity of a Real-time Algorithm for Temporal Gait Analysis using Inertial Measurement Units. *J Biomech*, *55*, 27-33. https://doi.org/10.1016/j.jbiomech.2017.02.016
- Aminian, K., Najafi, B., Büla, C., Leyvraz, P. F., & Robert, P. (2002). Spatio-temporal parameters of gait measured by an ambulatory system using miniature gyroscopes. *J Biomech*, *35*(5), 689-699. https://doi.org/10.1016/S0021-9290(02)00008-8
- Aminian, K., Rezakhanlou, K., De Andres, E., Fritsch, C., Leyvraz, P. F., & Robert, P. (1999). Temporal feature estimation during walking using miniature accelerometers: An analysis of gait improvement after hip arthroplasty. *Med Biol Eng Comput*, *37*(6), 686-691. https://doi.org/10.1007/BF02513368
- Aminian, K., Robert, P., Jequier, E., & Schutz, Y. (1995). Estimation of speed and incline of walking using neural network. *IEEE transactions on instrumentation and measurement*, *44*(3), 743-746. https://doi.org/10.1109/19.387322

- Anwary, A. R., Yu, H., & Vassallo, M. (2018). An Automatic gait feature extraction method for identifying gait asymmetry using wearable sensors. *Sensors (Basel)*, *18*(2), 676. https://doi.org/10.3390/s18020676
- Atallah, L., Wiik, A., Lo, B., Cobb, J. P., Amis, A. A., & Yang, G. Z. (2014). Gait asymmetry detection in older adults using a light ear-worn sensor. *Physiol. Meas*, *35*(5), N29-N40. https://doi.org/10.1088/0967-3334/35/5/N29
- Barnett, A., & Cerin, E. (2006). Individual calibration for estimating free-living walking speed using the MTI monitor. *Med Sci Sports Exerc*, *38*(4), 761-767. https://doi.org/10.1249/01.mss.0000210206.55941.b2
- Bejarano, N. C., Ambrosini, E., Pedrocchi, A., Ferrigno, G., Monticone, M., & Ferrante, S. (2014). A novel adaptive, real-time algorithm to detect gait events from wearable sensors. *IEEE transactions on neural systems and rehabilitation engineering*, *23*(3), 413-422.
- Ben Mansour, K., Rezzoug, N., & Gorce, P. (2015). Foot side detection from lower lumbar spine acceleration. *Gait Posture*, *42*(3), 386-389. https://doi.org/10.1016/j.gaitpost.2015.05.021
- Bertuletti, S., Della Croce, U., & Cereatti, A. (2019). A wearable solution for accurate step detection based on the direct measurement of the inter-foot distance. *J Biomech*, *84*, 274-277. https://doi.org/10.1016/j.jbiomech.2018.12.039
- Brodie, M. A. D., Coppens, M. J. M., Lord, S. R., Lovell, N. H., Gschwind, Y. J., Redmond, S. J., . . . Delbaere, K. (2016). Wearable pendant device monitoring using new wavelet-based methods shows daily life and laboratory gaits are different. *Med Biol Eng Comput*, *54*(4), 663-674. https://doi.org/10.1007/s11517-015-1357-9
- Bugané, F., Benedetti, M. G., Casadio, G., Attala, S., Biagi, F., Manca, M., & Leardini, A. (2012). Estimation of spatial-temporal gait parameters in level walking based on a single accelerometer: Validation on normal subjects by standard gait analysis. *Comput Methods Programs Biomed*, *108*(1), 129-137. https://doi.org/10.1016/j.cmpb.2012.02.003

- Byun, S., Han, J. W., Kim, T. H., & Kim, K. W. (2016). Test-retest reliability and concurrent validity of a single tri-axial accelerometer-based gait analysis in older adults with normal cognition. *PLoS One*, *11*(7), e0158956-e0158956. https://doi.org/10.1371/journal.pone.0158956
- Bötzel, K., Olivares, A., Cunha, J. P., Górriz Sáez, J. M., Weiss, R., & Plate, A. (2018). Quantification of gait parameters with inertial sensors and inverse kinematics. *J Biomech*, *72*, 207-214. https://doi.org/10.1016/j.jbiomech.2018.03.012
- Cho, Y.-S., Jang, S.-H., Cho, J.-S., Kim, M.-J., Lee, H. D., Lee, S. Y., & Moon, S.-B. (2018). Evaluation of validity and reliability of inertial measurement unit-based gait analysis systems. *Ann Rehabil Med*, *42*(6), 872-883. https://doi.org/10.5535/arm.2018.42.6.872
- Cole, M. H., Van Den Hoorn, W., Kavanagh, J. K., Morrison, S., Hodges, P. W., Smeathers, J. E., & Kerr, G. K. (2014). Concurrent validity of accelerations measured using a tri-axial inertial measurement unit while walking on firm, compliant and uneven surfaces. *PLoS One*, *9*(5), e98395-e98395. https://doi.org/10.1371/journal.pone.0098395
- Das, R., Hooda, N., & Kumar, N. (2019). A Novel Approach for Real-Time Gait Events Detection Using Developed Wireless Foot Sensor Module. *IEEE sensors letters*, *3*(6), 1-4. https://doi.org/10.1109/LSENS.2019.2914719
- De Vroey, H., Staes, F., Weygers, I., Vereecke, E., Vanrenterghem, J., Deklerck, J., . . . Claeys, K. (2018). The implementation of inertial sensors for the assessment of temporal parameters of gait in the knee arthroplasty population. *Clin Biomech (Bristol, Avon)*, *54*, 22-27. https://doi.org/10.1016/j.clinbiomech.2018.03.002
- Del Din, S., Godfrey, A., & Rochester, L. (2016). Validation of an Accelerometer to Quantify a Comprehensive Battery of Gait Characteristics in Healthy Older Adults and Parkinson's Disease: Toward Clinical and at Home Use. *IEEE J Biomed Health Inform*, *20*(3), 838-847. https://doi.org/10.1109/JBHI.2015.2419317
- Del Din, S., Hickey, A., Hurwitz, N., Mathers, J. C., Rochester, L., & Godfrey, A. (2016). Measuring gait with an accelerometer-based wearable: influence of device location, testing protocol and age. *Physiol. Meas*, *37*(10), 1785-1797. https://doi.org/10.1088/0967-3334/37/10/1785

- Derungs, A., Schuster-Amft, C., & Amft, O. (2018). Longitudinal walking analysis in hemiparetic patients using wearable motion sensors: Is there convergence between body sides? *Front Bioeng Biotechnol*, *6*(MAY), 57-57. https://doi.org/10.3389/fbioe.2018.00057
- Dobkin, B. H., Xu, X., Batalin, M., Thomas, S., & Kaiser, W. (2011). Reliability and validity of bilateral ankle accelerometer algorithms for activity recognition and walking speed after stroke. *Stroke*, *42*(8), 2246-2250. https://doi.org/10.1161/STROKEAHA.110.611095
- Fasel, B., Duc, C., Dadashi, F., Bardyn, F., Savary, M., Farine, P.-A., & Aminian, K. (2017). A wrist sensor and algorithm to determine instantaneous walking cadence and speed in daily life walking. *Med Biol Eng Comput*, *55*(10), 1773-1785. https://doi.org/10.1007/s11517-017-1621-2
- Ferrari, A., Ginis, P., Hardegger, M., Casamassima, F., Rocchi, L., & Chiari, L. (2015). A mobile Kalman-filter based solution for the real-time estimation of spatio-temporal gait parameters. *IEEE transactions on neural systems and rehabilitation engineering*, *24*(7), 764-773.
- Figueiredo, J., Felix, P., Costa, L., Moreno, J. C., & Santos, C. P. (2018). Gait Event Detection in Controlled and Real-Life Situations: Repeated Measures From Healthy Subjects. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *26*(10), 1945-1956. https://doi.org/10.1109/tnsre.2018.2868094
- Flood, M. W., O'Callaghan, B. P. F., & Lowery, M. M. (2020). Gait Event Detection From Accelerometry Using the Teager-Kaiser Energy Operator. *IEEE Trans Biomed Eng*, *67*(3), 658-666. https://doi.org/10.1109/TBME.2019.2919394
- Fourati, H. (2015). Heterogeneous Data Fusion Algorithm for Pedestrian Navigation via Foot-Mounted Inertial Measurement Unit and Complementary Filter. *IEEE transactions on instrumentation and measurement*, *64*(1), 221-229. https://doi.org/10.1109/TIM.2014.2335912

- Fusca, M., Perego, P., & Andreoni, G. (2018). Method for Wearable Kinematic Gait Analysis Using a Harmonic Oscillator Applied to the Center of Mass. *Journal of sensors*, *2018*, 1-14. https://doi.org/10.1155/2018/4548396
- Giladi, N., Sharifi, S., van Vugt, J. P. P., Hausdorff, J. M., Plotnik, M., & Weiss, A. (2011). Toward automated, at-home assessment of mobility among patients with Parkinson Disease, using a body-worn accelerometer. *Neurorehabil Neural Repair*, *25*(9), 810-818. https://doi.org/10.1177/1545968311424869
- Godfrey, A., Del Din, S., Barry, G., Mathers, J. C., & Rochester, L. (2015). Instrumenting gait with an accelerometer: A system and algorithm examination. *Med Eng Phys*, *37*(4), 400-407. https://doi.org/10.1016/j.medengphy.2015.02.003
- Gouwanda, D., & Gopalai, A. A. (2015). A robust real-time gait event detection using wireless gyroscope and its application on normal and altered gaits. *Med Eng Phys*, *37*(2), 219-225. https://doi.org/10.1016/j.medengphy.2014.12.004
- Greene, B. R., McGrath, D., O'Neill, R., O'Donovan, K. J., Burns, A., & Caulfield, B. (2010). An adaptive gyroscope-based algorithm for temporal gait analysis. *Medical & biological engineering & computing*, *48*(12), 1251-1260.
- Han, J., Jeon, H. S., Yi, W. J., Jeon, B. S., & Park, K. S. (2009). Adaptive windowing for gait phase discrimination in Parkinsonian gait using 3-axis acceleration signals. *Med Biol Eng Comput*, *47*(11), 1155-1164. https://doi.org/10.1007/s11517-009-0521-5
- Han, Y. C., Wong, K. I., & Murray, I. (2019a). Gait Phase Detection for Normal and Abnormal Gaits Using IMU. *IEEE sensors journal*, *19*(9), 3439-3448. https://doi.org/10.1109/JSEN.2019.2894143
- Han, Y. C., Wong, K. I., & Murray, I. (2019b). Stride Length Estimation Based on a Single Shank's Gyroscope. *IEEE sensors letters*, *3*(10), 1-4. https://doi.org/10.1109/LSENS.2019.2942177
- Hanlon, M., & Anderson, R. (2006). Real-time gait event detection using wearable sensors. *Gait & posture*, *24*, S127-S128. https://doi.org/10.1016/j.gaitpost.2006.11.088

- Hannink, J., Kautz, T., Pasluosta, C. F., Gasmann, K.-G., Klucken, J., & Eskofier, B. M. (2017). Sensor-Based Gait Parameter Extraction With Deep Convolutional Neural Networks. *IEEE J Biomed Health Inform*, *21*(1), 85-93. https://doi.org/10.1109/JBHI.2016.2636456
- Hao, M., Chen, K., & Fu, C. (2019). Smoother-Based 3-D Foot Trajectory Estimation Using Inertial Sensors. *IEEE Trans Biomed Eng*, *66*(12), 3534-3542. https://doi.org/10.1109/TBME.2019.2907322
- Hickey, A., Del Din, S., Rochester, L., & Godfrey, A. (2016). Detecting free-living steps and walking bouts: validating an algorithm for macro gait analysis. *Physiol. Meas*, *38*(1), N1-N15. https://doi.org/10.1088/1361-6579/38/1/N1
- Hsing-Cheng, C., Yu-Liang, H., Shih-Chin, Y., Jung-Chih, L., & Zhi-Hao, W. (2016). A Wearable Inertial Measurement System With Complementary Filter for Gait Analysis of Patients With Stroke or Parkinson's Disease. *IEEE access*, *4*, 8442-8453. https://doi.org/10.1109/ACCESS.2016.2633304
- Hu, J.-S., Sun, K.-C., & Cheng, C.-Y. (2013). A Kinematic Human-Walking Model for the Normal-Gait-Speed Estimation Using Tri-Axial Acceleration Signals at Waist Location. *IEEE Trans Biomed Eng*, *60*(8), 2271-2279. https://doi.org/10.1109/TBME.2013.2252345
- Hundza, S. R., Hook, W. R., Harris, C. R., Mahajan, S. V., Leslie, P. A., Spani, C. A., . . . Livingston, N. J. (2013). Accurate and reliable gait cycle detection in Parkinson's disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *22*(1), 127-137.
- Hwang, T.-H., Reh, J., Effenberg, A. O., & Blume, H. (2018). Real-Time Gait Analysis Using a Single Head-Worn Inertial Measurement Unit. *IEEE transactions on consumer electronics*, *64*(2), 240-248. https://doi.org/10.1109/TCE.2018.2843289
- Ichinoseki-Sekine, N., Kuwae, Y., Higashi, Y., Fujimoto, T., Sekine, M., & Tamura, T. (2006). Improving the accuracy of pedometer used by the elderly with the FFT algorithm. *Med Sci Sports Exerc*, *38*(9), 1674-1681. https://doi.org/10.1249/01.mss.0000227641.68360.c2

- Jarchi, D., Wong, C., Kwasnicki, R. M., Heller, B., Tew, G. A., & Yang, G.-Z. (2014). Gait Parameter Estimation From a Miniaturized Ear-Worn Sensor Using Singular Spectrum Analysis and Longest Common Subsequence. *IEEE Trans Biomed Eng*, *61*(4), 1261-1273. https://doi.org/10.1109/TBME.2014.2299772
- Ji, Q., Yang, L., Li, W., Zhou, C., & Ye, X. (2018). Real-time gait event detection in a real-world environment using a laser-ranging sensor and gyroscope fusion method. *Physiol. Meas*, *39*(12), 125003-125003. https://doi.org/10.1088/1361-6579/aae7ee
- Kawamura, K., Morita, Y., Okamoto, J., Saito, K., Sessa, S., Zecca, M., . . . Fujie, M. G. (2012). Gait phase detection using foot acceleration for estimating ground reaction force in long distance gait rehabilitation. *Journal of Robotics and Mechatronics*, *24*(5), 828-837.
- Köse, A., Cereatti, A., & Della Croce, U. (2012). Bilateral step length estimation using a single inertial measurement unit attached to the pelvis. *J Neuroeng Rehabil*, *9*(1), 9-9. https://doi.org/10.1186/1743-0003-9-9
- Lee, J.-A., Cho, S.-H., Lee, Y.-J., Yang, H.-K., & Lee, J.-W. (2010). Portable Activity Monitoring System for Temporal Parameters of Gait Cycles. *J Med Syst*, *34*(5), 959-966. https://doi.org/10.1007/s10916-009-9311-8
- Lee, J. K., & Park, E. J. (2011). Quasi real-time gait event detection using shank-attached gyroscopes. *Med Biol Eng Comput*, *49*(6), 707-712. https://doi.org/10.1007/s11517-011-0736-0
- Lee, M., Youm, C., Jeon, J., Cheon, S.-M., & Park, H. (2018). Validity of shoe-type inertial measurement units for Parkinson's disease patients during treadmill walking. *J Neuroeng Rehabil*, *15*(1), 38-38. https://doi.org/10.1186/s12984-018-0384-9
- Li, Q., Young, M., Naing, V., & Donelan, J. M. (2010). Walking speed estimation using a shank-mounted inertial measurement unit. *J Biomech*, *43*(8), 1640-1643. https://doi.org/10.1016/j.jbiomech.2010.01.031
- López-Nava, I. H., Muñoz-Meléndez, A., Pérez Sanpablo, A. I., Alessi Montero, A., Quiñones Urióstegui, I., & Núñez Carrera, L. (2016). Estimation of temporal gait

parameters using Bayesian models on acceleration signals. *Comput Methods Biomech Biomed Engin*, *19*(4), 396-403. https://doi.org/10.1080/10255842.2015.1032945

- Mannini, A., Genovese, V., & Sabatin, A. M. (2014). Online Decoding of Hidden Markov Models for Gait Event Detection Using Foot-Mounted Gyroscopes. *IEEE J Biomed Health Inform*, *18*(4), 1122-1130. https://doi.org/10.1109/JBHI.2013.2293887
- Mannini, A., & Sabatini, A. M. (2014). Walking speed estimation using foot-mounted inertial sensors: Comparing machine learning and strap-down integration methods. *Med Eng Phys*, *36*(10), 1312-1321. https://doi.org/10.1016/j.medengphy.2014.07.022
- Maqbool, H. F., Husman, M. A. B., Awad, M. I., Abouhossein, A., & Dehghani-Sanij, A. A. (2015). Real-time gait event detection for transfemoral amputees during ramp ascending and descending. *EMBC*, *2015-*, 4785-4788. https://doi.org/10.1109/EMBC.2015.7319464
- Mariani, B., Hoskovec, C., Rochat, S., Büla, C., Penders, J., & Aminian, K. (2010). 3D gait assessment in young and elderly subjects using foot-worn inertial sensors. *J Biomech*, *43*(15), 2999-3006. https://doi.org/10.1016/j.jbiomech.2010.07.003
- Mariani, B., Jiménez, M. C., Vingerhoets, F. J. G., & Aminian, K. (2013). On-Shoe Wearable Sensors for Gait and Turning Assessment of Patients With Parkinson's Disease. *IEEE Trans Biomed Eng*, *60*(1), 155-158. https://doi.org/10.1109/TBME.2012.2227317
- Martindale, C. F., Roth, N., Gassner, H., List, J., Regensburger, M., Eskofier, B. M., & Kohl, Z. (2020). Technical Validation of an Automated Mobile Gait Analysis System for Hereditary Spastic Paraplegia Patients. *IEEE J Biomed Health Inform*, *24*(5), 1490-1499. https://doi.org/10.1109/JBHI.2019.2937574
- Martinez-Hernandez, U., & Dehghani-Sanij, A. A. (2018). Adaptive Bayesian inference system for recognition of walking activities and prediction of gait events using wearable sensors. *Neural Networks*, *102*, 107-119.
- McGinnis, R. S., Mahadevan, N., Moon, Y., Seagers, K., Sheth, N., Wright, J. A., . . . Patel, S. (2017). A machine learning approach for gait speed estimation using skin-mounted

wearable sensors: From healthy controls to individuals with multiple sclerosis. *PLoS One*, *12*(6), e0178366-e0178366. https://doi.org/10.1371/journal.pone.0178366

- Mico-Amigo, M. E., Kingma, I., Ainsworth, E., Walgaard, S., Niessen, M., van Lummel, R. C., & van Dieen, J. H. (2016). A novel accelerometry-based algorithm for the detection of step durations over short episodes of gait in healthy elderly. *J Neuroeng Rehabil*, *13*(1), 38-38. https://doi.org/10.1186/s12984-016-0145-6
- Misu, S., Asai, T., Ono, R., Sawa, R., Tsutsumimoto, K., Ando, H., & Doi, T. (2017). Development and validity of methods for the estimation of temporal gait parameters from heel-attached inertial sensors in younger and older adults. *Gait Posture*, *57*, 295-298. https://doi.org/10.1016/j.gaitpost.2017.06.022
- Miyazaki, S. (1997). Long-term unrestrained measurement of stride length and walking velocity utilizing a piezoelectric gyroscope. *IEEE Trans Biomed Eng*, *44*(8), 753-759. https://doi.org/10.1109/10.605434
- Motoi, K., Taniguchi, S., Baek, M., Wakugawa, M., Sonoda, T., Yuji, T., . . . Yamakoshi, K.-I. (2012). Development of a wearable gait monitoring system for evaluating efficacy of walking training in rehabilitation. *Sensors and materials*, *24*(6), 359-373. https://doi.org/10.18494/SAM.2012.792
- Nguyen, A., Roth, N., Ghassemi, N. H., Hannink, J., Seel, T., Klucken, J., . . . Eskofier, B. M. (2019). Development and clinical validation of inertial sensor-based gait-clustering methods in Parkinson's disease. *J Neuroeng Rehabil*, *16*(1), 77-77. https://doi.org/10.1186/s12984-019-0548-2
- O'Brien, M. K., Hidalgo-Araya, M. D., Mummidisetty, C. K., Vallery, H., Ghaffari, R., Rogers, J. A., . . . Jayaraman, A. (2019). Augmenting Clinical Outcome Measures of Gait and Balance with a Single Inertial Sensor in Age-Ranged Healthy Adults. *Sensors (Basel)*, *19*(20), 4537. https://doi.org/10.3390/s19204537
- Papi, E., Osei-Kuffour, D., Chen, Y.-M. A., & McGregor, A. H. (2015). Use of wearable technology for performance assessment: A validation study. *Med Eng Phys*, *37*(7), 698-704. https://doi.org/10.1016/j.medengphy.2015.03.017

- Pappas, I. P., Popovic, M. R., Keller, T., Dietz, V., & Morari, M. (2001). A reliable gait phase detection system. *IEEE Transactions on neural systems and rehabilitation engineering*, *9*(2), 113-125.
- Perrin, O., Terrier, P., Ladetto, Q., Merminod, B., & Schutz, Y. (2000). Improvement of walking speed prediction by accelerometry and altimetry, validated by satellite positioning. *Med Biol Eng Comput*, *38*(2), 164-168. https://doi.org/10.1007/BF02344771
- Pham, M. H., Elshehabi, M., Haertner, L., Din, S. D., Srulijes, K., Heger, T., . . . Maetzler, W. (2017). 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*, *8*(SEP), 457-457. https://doi.org/10.3389/fneur.2017.00457
- Prateek, G. V., Mazzoni, P., Earhart, G. M., & Nehorai, A. (2020). Gait Cycle Validation and Segmentation Using Inertial Sensors. *IEEE Trans Biomed Eng*, *67*(8), 2132-2144. https://doi.org/10.1109/TBME.2019.2955423
- Punt, M., van Alphen, B., van de Port, I. G., van Dieen, J. H., Michael, K., Outermans, J. C., & Wittink, H. (2014). Clinimetric properties of a novel feedback device for assessing gait parameters in stroke survivors. *J Neuroeng Rehabil*, *11*(1), 30-30. https://doi.org/10.1186/1743-0003-11-30
- Qiu, S., Liu, L., Wang, Z., Li, S., Zhao, H., Wang, J., . . . Tang, K. (2019). Body sensor network-based gait quality assessment for clinical decision-support via multi-sensor fusion. *Ieee Access*, *7*, 59884-59894.
- Qiu, S., Wang, Z., Zhao, H., & Hu, H. (2016). Using Distributed Wearable Sensors to Measure and Evaluate Human Lower Limb Motions. *IEEE transactions on instrumentation and measurement*, *65*(4), 939-950. https://doi.org/10.1109/TIM.2015.2504078
- Rampp, A., Barth, J., Schuelein, S., Gassmann, K.-G., Klucken, J., & Eskofier, B. M. (2015). Inertial Sensor-Based Stride Parameter Calculation From Gait Sequences in Geriatric

Patients. *IEEE Trans Biomed Eng*, *62*(4), 1089-1097. https://doi.org/10.1109/TBME.2014.2368211

- Rebula, J. R., Ojeda, L. V., Adamczyk, P. G., & Kuo, A. D. (2013). Measurement of foot placement and its variability with inertial sensors. *Gait Posture*, *38*(4), 974-980. https://doi.org/10.1016/j.gaitpost.2013.05.012
- Sabatini, A. M., Martelloni, C., Scapellato, S., & Cavallo, F. (2005). Assessment of walking features from foot inertial sensing. *IEEE Trans Biomed Eng*, *52*(3), 486-494. https://doi.org/10.1109/TBME.2004.840727
- Salarian, A., Russmann, H., Vingerhoets, F. J. G., Dehollain, C., Blanc, Y., Burkhard, P. R., & Aminian, K. (2004). Gait assessment in Parkinson's disease: toward an ambulatory system for long-term monitoring. *IEEE Trans Biomed Eng*, *51*(8), 1434-1443. https://doi.org/10.1109/TBME.2004.827933
- Sejdic, E., Lowry, K. A., Bellanca, J., Perera, S., Redfern, M. S., & Brach, J. S. (2016). Extraction of Stride Events From Gait Accelerometry During Treadmill Walking. *IEEE J Transl Eng Health Med*, *4*, 1-11. https://doi.org/10.1109/JTEHM.2015.2504961
- Selles, R. W., Formanoy, M. A., Bussmann, J. B., Janssens, P. J., & Stam, H. J. (2005). Automated estimation of initial and terminal contact timing using accelerometers; development and validation in transtibial amputees and controls. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *13*(1), 81-88.
- Soaz, C., & Diepold, K. (2016). Step Detection and Parameterization for Gait Assessment Using a Single Waist-Worn Accelerometer. *IEEE Trans Biomed Eng*, *63*(5), 933-942. https://doi.org/10.1109/TBME.2015.2480296
- Soltani, A., Dejnabadi, H., Savary, M., & Aminian, K. (2020). Real-World Gait Speed Estimation Using Wrist Sensor: A Personalized Approach. *IEEE J Biomed Health Inform*, *24*(3), 658-668. https://doi.org/10.1109/JBHI.2019.2914940
- Song, M., & Kim, J. (2018). An Ambulatory Gait Monitoring System with Activity Classification and Gait Parameter Calculation Based on a Single Foot Inertial Sensor. *IEEE Trans Biomed Eng*, *65*(4), 885-893. https://doi.org/10.1109/TBME.2017.2724543

- Sprager, S., & Juric, M. B. (2018). Robust stride segmentation of inertial signals based on local cyclicity estimation. *Sensors (Basel)*, *18*(4), 1091. https://doi.org/10.3390/s18041091
- Storm, F. A., Buckley, C. J., & Mazzà, C. (2016). Gait event detection in laboratory and real life settings: Accuracy of ankle and waist sensor based methods. *Gait Posture*, *50*, 42-46. https://doi.org/10.1016/j.gaitpost.2016.08.012
- Storm, F. A., Nair, K. P. S., Clarke, A. J., Van der Meulen, J. M., & Mazzà, C. (2018). Free-living and laboratory gait characteristics in patients with multiple sclerosis. *PLoS One*, *13*(5), e0196463-e0196463. https://doi.org/10.1371/journal.pone.0196463
- Supratak, A., Datta, G., Gafson, A. R., Nicholas, R., Guo, Y., & Matthews, P. M. (2018). Remote monitoring in the home validates clinical gait measures for multiple sclerosis. *Front Neurol*, *9*, 561-561. https://doi.org/10.3389/fneur.2018.00561
- Tao, S., Zhang, X., Cai, H., Lv, Z., Hu, C., & Xie, H. (2018). Gait based biometric personal authentication by using MEMS inertial sensors. *Journal of ambient intelligence and humanized computing*, *9*(5), 1705-1712. https://doi.org/10.1007/s12652-018-0880-6
- Teufl, W., Lorenz, M., Miezal, M., Taetz, B., Fröhlich, M., & Bleser, G. (2018). Towards inertial sensor based mobile gait analysis: Event-detection and spatio-temporal parameters. *Sensors (Basel)*, *19*(1), 38. https://doi.org/10.3390/s19010038
- van Oeveren, B. T., de Ruiter, C. J., Beek, P. J., Rispens, S. M., & van Dieën, J. H. (2018). An adaptive, real-time cadence algorithm for unconstrained sensor placement. *Med Eng Phys*, *52*(February), 49-58. https://doi.org/10.1016/j.medengphy.2017.12.007
- Vincenzo, G., Andrea, M., & Angelo, M. S. (2017). A Smartwatch Step Counter for Slow and Intermittent Ambulation. *IEEE access*, *5*, 13028-13037. https://doi.org/10.1109/ACCESS.2017.2702066
- Wang, L., Sun, Y., Li, Q., & Liu, T. (2018). Estimation of Step Length and Gait Asymmetry Using Wearable Inertial Sensors. *IEEE sensors journal*, *18*(9), 3844-3851. https://doi.org/10.1109/JSEN.2018.2815700

- Wang, Z., Zhao, H., Qiu, S., & Gao, Q. (2015). Stance-Phase Detection for ZUPT-Aided Foot-Mounted Pedestrian Navigation System. *IEEE/ASME transactions on mechatronics*, *20*(6), 3170-3181. https://doi.org/10.1109/TMECH.2015.2430357
- Yang, C.-C., Hsu, Y.-L., Shih, K.-S., & Lu, J.-M. (2011). Real-time gait cycle parameter recognition using a wearable accelerometry system. *Sensors (Basel)*, *11*(8), 7314-7326. https://doi.org/10.3390/s110807314
- Yang, S., & Li, Q. (2012). IMU-based ambulatory walking speed estimation in constrained treadmill and overground walking. *Comput Methods Biomech Biomed Engin*, *15*(3), 313-322. https://doi.org/10.1080/10255842.2010.534465
- Yang, S., Zhang, J.-T., Novak, A. C., Brouwer, B., & Li, Q. (2012). Estimation of spatio-temporal parameters for post-stroke hemiparetic gait using inertial sensors. *Gait Posture*, *37*(3), 354-358. https://doi.org/10.1016/j.gaitpost.2012.07.032
- Yuwono, M., Su, S. W., Guo, Y., Moulton, B. D., & Nguyen, H. T. (2014). Unsupervised nonparametric method for gait analysis using a waist-worn inertial sensor. *Applied soft computing*, *14*, 72-80. https://doi.org/10.1016/j.asoc.2013.07.027
- Zhang, H., Guo, Y., & Zanotto, D. (2019). Accurate ambulatory gait analysis in walking and running using machine learning models. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *28*(1), 191-202.
- Zhang, H., Zhang, J., Zhou, D., Wang, W., Li, J., Ran, F., & Ji, Y. (2015). Axis-Exchanged Compensation and Gait Parameters Analysis for High Accuracy Indoor Pedestrian Dead Reckoning. *Journal of sensors*, *2015*, 1-13. https://doi.org/10.1155/2015/915837
- Zhang, Y., Li, Y., Peng, C., Mou, D., Li, M., & Wang, W. (2018). The height-adaptive parameterized step length measurement method and experiment based on motion parameters. *Sensors (Basel)*, *18*(4), 1039. https://doi.org/10.3390/s18041039
- Zhao, H., Wang, Z., Qiu, S., Shen, Y., Zhang, L., Tang, K., & Fortino, G. (2018). Heading drift reduction for foot-mounted inertial navigation system via multi-sensor fusion and dual-gait analysis. *IEEE Sensors Journal*, *19*(19), 8514-8521.

- Zhao, H., Wang, Z., Qiu, S., Wang, J., Xu, F., Wang, Z., & Shen, Y. (2019). Adaptive gait detection based on foot-mounted inertial sensors and multi-sensor fusion. *Information fusion*, *52*, 157-166. https://doi.org/10.1016/j.inffus.2019.03.002
- Zhou, H., Ji, N., Samuel, O. W., Cao, Y., Zhao, Z., Chen, S., & Li, G. (2016). Towards real-time detection of gait events on different terrains using time-frequency analysis and peak heuristics algorithm. *Sensors (Basel)*, *16*(10), 1634-1634. https://doi.org/10.3390/s16101634
- Zihajehzadeh, S., & Park, E. J. (2016). Regression model-based walking speed estimation using wrist-worn inertial sensor. *PLoS One*, *11*(10), e0165211-e0165211. https://doi.org/10.1371/journal.pone.0165211
- Zijlstra, W., & Hof, A. L. (2003). Assessment of spatio-temporal gait parameters from trunk accelerations during human walking. *Gait Posture*, *18*(2), 1-10. https://doi.org/10.1016/S0966-6362(02)00190-X

Appendix Two: Acknowledgement of collaborative work within the thesis

Acknowledgement of collaborative work within the thesis

The candidate confirms that the work submitted is their own, except where work that has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

Publications included in thesis:

Scott K, Bonci T, Alcock L, Buckley E, Hansen C, Gazit E, Schwickert L, Cereatti A, Mazzà C, on behalf of the Mobilise-D Consortium. A Quality Control Check to Ensure Comparability of Stereophotogrammetric Data between Sessions and Systems. Sensors. 2021; 21(24):8223. doi: 10.3390/s21248223

Author Contributions: Conceptualization, K.S., T.B., L.A., A.C. and C.M.; methodology, K.S., T.B., L.A., A.C. and C.M.; software, K.S., T.B., L.A., E.B, C.H., E.G, L.S., A.C. and C.M.; validation, K.S., A.C., and C.M.; formal analysis, K.S.; investigation, K.S., T.B., L.A., E.B., C.H., E.G., L.S., A.C. and C.M.; resources, K.S. and C.M.; data curation, K.S.; writing – original draft preparation, K.S. and C.M.; writing – review and editing, K.S., T.B., L.A., E.B., C.H., E.G., L.S., A.C. and C.M.; visualization, K.S., A.C. and C.M.; supervision, K.S., A.C. and C.M.; project administration, K.S.; funding acquisition, C.M. All authors have read and agreed to the published version of the manuscript.

Scott, Kirsty, Bonci Tecla, Salis Francesca, Bertuletti Stefano, Caruso Marco, Alcock Lisa, Hansen Clint, Buckley Ellen, Schwickert Lars, Gazit Eran, Cereatti Andrea, Lorenzo Chiari, Sharrack Basil, Walter Maetzler, Becker Clemens, Hausdorff Jeffrey, Vogiatzis, Ioannis, Brown Philip, Del Din Silvia, Eskofier Björn, Helbostad Jorunn, Paraschiv-Ionescu Anisoara, Keogh Alison, Kirk Cameron, Kluge Felix, Micó-Amigo M. Encarna, Mueller Arne, Neatrour Isabel, Niessen Martjin, Palmerini Luca, Sillen Henrik, Singleton David, Ullrich Martin, Vereijken Beatrix, Yarnall Alison, Rochester Lynn and Mazzà Claudia. Design and validation of a multi-task, multi-context protocol for real-world gait simulation. J NeuroEngineering Rehabil 19, 141 (2022). doi: 10.1186/s12984-022-01116-1

Author Contributions: Study design: KS, CM, LR, AC, SDD, AM, BC, LC, VB. DATA COLLECTION: KS, TB, LA, EB, EG, CH, LS, PB, AK, GB. Design and implementation of experimental tools: KS, FS, SB, MC, AC, AK, MN, DS, LC, CM. Ethical approval, patient recruitment and clinical oversight: BS, WM, CB, JH, IV, IN, AY, LR. Data processing: KS, TB, SDD, FS, KA, BS, MC, LP, BE, AP, KC, FK, DS, MU, AK, EM. Statistical analysis: KS, AEC, SK, JGA, HS, SDD, AM. Manuscript initial drafting: KS, CM, WM, LC, LA. Figures and tables preparation: KS, TB, LA, SDD, CM. Manuscript revision: all co-authors. All authors read and approved the final manuscript.

.

Appendix Three: Ethics approval for pilot data collection

The University Of Sheffield.

Downloaded: 02/02/2023
Approved: 01/07/2019

Kirsty Scott
Registration number: 180300189
Mechanical Engineering
Programme: Mechanical Engineering

Dear Kirsty

PROJECT TITLE: Validation of an experimental protocol to determine the accuracy of a single inertial sensor system for assessing mobility in the real world
APPLICATION: Reference Number 029143

On behalf of the University ethics reviewers who reviewed your project, I am pleased to inform you that on 01/07/2019 the above-named project was **approved** on ethics grounds, on the basis that you will adhere to the following documentation that you submitted for ethics review:

- University research ethics application form 029143 (form submission date: 27/06/2019); (expected project end date: 21/06/2022).
- Participant information sheet 1065084 version 2 (27/06/2019).
- Participant consent form 1065085 version 1 (21/06/2019).

If during the course of the project you need to deviate significantly from the above-approved documentation please inform me since written approval will be required.

Your responsibilities in delivering this research project are set out at the end of this letter.

Yours sincerely

Jennifer Rowson
Ethics Administrator
Mechanical Engineering

Please note the following responsibilities of the researcher in delivering the research project:

- The project must abide by the University's Research Ethics Policy:https://www.sheffield.ac.uk/research-services/ethics-integrity/policy
- The project must abide by the University's Good Research & Innovation Practices Policy: https://www.sheffield.ac.uk/polopoly_fs/1.671066!/file/GRIPPolicy.pdf
- The researcher must inform their supervisor (in the case of a student) or Ethics Administrator (in the case of a member of staff) of any significant changes to the project or the approved documentation.
- The researcher must comply with the requirements of the law and relevant guidelines relating to security and confidentiality of personal data.
- The researcher is responsible for effectively managing the data collected both during and after the end of the project in line with best practice, and any relevant legislative, regulatory or contractual requirements.

Appendix Four: Abstract Published at ISPGR 2022 Conference

Abstract title: Inter-rater reliability when digitising the shoe perimeter; A personalised method for evaluating foot-ground interactions accounting for shoe shape

Kirsty Scott [1], Naomi Davis [1], Brook Galna [2], Lynn Rochester [3], Claudia Mazza [1] & Lisa Alcock [3]

[1] Department of Mechanical Engineering & INSIGNEO Institute for in silico Medicine, The University of Sheffield, Sheffield, UK

[2] Health Futures Institute, Murdoch University, Perth, Australia

[3] Institute of Translational and Clinical Research, Newcastle University, Newcastle upon Tyne, UK

Background & Aim

Foot-ground interactions (i.e. identifying gait events; heel strike/toe-off, foot clearance) are often measured during clinical gait evaluations using 3-dimensional stereophogrammetric systems. Movement of reflective markers affixed to the shoe are evaluated as trajectories reflecting a single location on the shoe (i.e. toe/heel;[1]) or may form part of a geometric model defining the foot segment[2]. Relying on a single marker for the detection of gait events or subsequent calculations of gait outcomes does not account for the variety of shoe shapes. Previous methods for digitising the foot have not investigated inter-rater reliability[3-5]. We propose a method for digitising the perimeter of the shoe that accommodates variation in shoe size, shoe shape and sole wear and evaluate inter-rater reliability.

Methods

Data was collected using a 10 camera motion capture system (Vicon MT160, 100Hz). A quality control check (QC check) quantified systematic error from the calibration[6]. Two 14mm reflective markers were affixed to the shoe over the calcaneus and distal toe. A triad of 9mm markers were affixed to the lateral aspect of the shoe for tracking. A custom digitiser (length 460mm) was constructed with four 9mm reflective markers configured to create a 3D local co-ordinate system. Three raters performed the foot digitisation procedure (novice, experienced, expert) on the same Converse shoe (size 5). Three trials per rater were completed. The shoe was suspended in the air and orientated so that the toes were pointing up. The digitiser point was used to guide tracing around the perimeter of the shoe (clockwise starting at the toe) and a virtual point was constructed (Matlab). The highest and lowest vertical point of the shoe were identified using peak detection and the forefoot/rearfoot were segmented. Virtual markers were distributed around the forefoot and rearfoot perimeter (Figure 1; inter-marker distance 6mm). The distance between the highest and lowest point of the shoe was quantified from the virtual markers and using a tape measure.

Results

The QC check indicated errors of 0.6mm and 0.2degrees for the dynamic calibration trial. Although the time to complete the foot calibration varied between rater, each rater performed the calibration successfully with the level of error only deviating in 1mm within rater and between rater, which is comparable to the systematic errors quantified from the system. The shoe length when measured manually was 261mm. The same metric extracted from the foot digitisation procedure was within 0.4% (mean of 260mm).

Conclusions

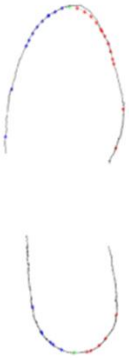
Digitising the perimeter of the foot can reliably be completed. Analyses to evaluate the influence of marker location on other outcomes, such as gait events and foot clearance, is underway.

[1] Alcock 2018 J Biomech [2] Alcock 2013 Gait & Posture [3] Startzell & Cavanagh, 1999 Hum Mov Sci [4] Loverro 2013 J Biomech [5] Telonio 2013 J Biomech [6] Scott 2021 Sensors

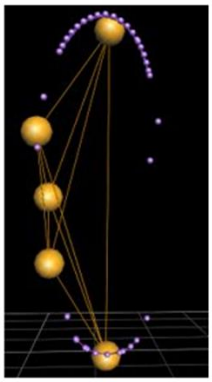
Digitiser

Converse Shoe Size 5

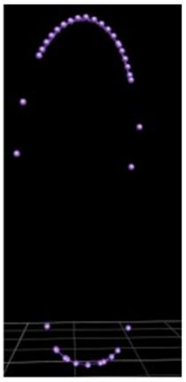
Trajectory of the virtual point (blue: forefoot, grey: midfoot, green: hindfoot)



Virtual foot markers in Matlab (inter-marker distance 6mm)



Distal toe and calcaneus markers, triad of tracking markers (lateral foot)



Virtual foot markers in Vicon (inter-marker distance 6mm)