



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN
SCIENZE E TECNOLOGIE DELLA SALUTE

Ciclo 36

Settore Concorsuale: 09/G2 - BIOINGEGNERIA

Settore Scientifico Disciplinare: ING-IND/34 – BIOINGEGNERIA INDUSTRIALE

CREDIBILITY OF DIGITAL HEALTH PREDICTORS OF HUMAN MOVEMENT

Presentata da: Francesca Bottin

Coordinatore Dottorato

Igor Diemberger

Supervisore

Marco Viceconti

Co-Supervisor

Lynn Rochester

Giorgio Davico

Esame finale anno 2024

Abstract

Human mobility, a vital sign of health, involves intricate coordination among the physiological systems. Any system disruptions likely result in reduced mobility, in turn leading to adverse clinical outcomes, such as the loss of independence and mortality. The consequences of the loss of mobility extend beyond personal well-being to societal and healthcare implications, calling for collaborative efforts to develop innovative solutions. Novel technologies such as wearable sensors and computational simulations have revolutionised the understanding of human mobility, offering unprecedented insights into human biomechanics. The availability of miniaturised wearable sensors, and advanced data processing algorithms, enable the monitoring of real-world mobility over long periods of time and to extract various parameters of clinical interest (e.g., cadence, gait asymmetry), also in pathological populations exhibiting atypical or slow gait. Concurrently, advancements in the field of *in silico* medicine have shifted the paradigm in the analysis of human biomechanics. Computer models, especially musculoskeletal dynamics models, can predict the human body's behaviour, offering various potential applications in clinical settings, such as personalised treatment or surgical plans. Understanding the limitations and potential errors of these technologies is crucial. Moreover, before any such technologies can be marketed or used in the clinics, their credibility must be established. This process requires rigorous testing procedures (against established gold standards) and remains complex, time-consuming, and occasionally expensive and ambiguous. In this context, this PhD thesis aimed to investigate and (eventually) perform the credibility assessment of two models: analytics software for wearable sensor data and (neuro)musculoskeletal dynamics models for identifying the primary cause for the loss of muscle force (i.e., dynapenia), mirroring the two distinct projects (i.e., the Mobilise-D and the ForceLoss projects, respectively). Both projects focused on conditions that negatively impact human movement (Parkinson's disease, multiple sclerosis, chronic obstructive pulmonary disease and proximal femur fracture in Mobilise-D, and the loss of muscle force in ForceLoss), and shared the need for the identification of a process to perform the credibility assessment for the respective solutions, intended for drug development and clinical decision-making. The Mobilise-D project, EU-funded, focused on the use of mobility-related parameters extracted by analytics software from continuous wearable sensor data as a new mobility biomarker for drug development. To deeply understand the regulatory processes needed for the qualification of this new methodology, engagements with the regulatory authorities (i.e., the European Medicine Agency – EMA – and the Food and Drug Administration – FDA) have occurred. Positive and constructive feedback were received in response to two requests for qualification advice to the EMA on the use of real-world mobility to generate evidence for inclusion in the marketing authorisation of new

medicines and from an interaction with the FDA CDRH Division through which we sought advice on the process to achieve the regulatory qualification (certification) of the Mobilise-D analytics software as a Software as Medical Device. On the other hand, the ForceLoss project delved into the use of subject-specific musculoskeletal models to support the differential diagnosis of dynapenia. A new framework was designed by combining both experimental measurements and computer models and simulations to enable the identification of the primary causes of dynapenia. Before applying the pipeline on twenty osteoarthritic patients candidates for primary total knee arthroplasty surgery, the framework was tested on a control group of healthy subjects.

In addition to the credibility assessment of a new methodology in a shared clinical context related to the loss of mobility, the overarching goal was to lower the barriers by sharing the acquired and processed experimental data as well as the insights derived from engagements with regulatory authorities.

List of the main Abbreviations

3D	Three-Dimensional
ADL	Activity of Daily Living
ASME	American Society of Mechanical Engineering
BQP	Biomarker Qualification Program
CBER	Center For Biologics Evaluation and Research
CDER	Center For Drug Evaluation and Research
CDRH	Center For Devices and Radiological Health
CEN	European Committee for Standardization (Comité Européen De Normalisation)
CHMP	Committee for Medicinal Products for Human Use
CNS	Central Nervous System
COPD	Chronic Obstructive Pulmonary Disease
CoU	Context of Use
CT	Computed Tomography
CVS	Clinical Validation Study
DDT	Drug Development Tools
DMO	Digital Mobility Outcome
DoF	Degree Of Freedom
DT	Digital Twin
EMA	European Medicine Agency
EMG	Electromyography
EU	European Union
FDA	Food & Drug Administration
HA	Healthy Adult
IMDRF	International Medical Device Regulators Forum
IMU	Inertial Measurement Unit
ISO	International Organization for Standardization
LLFDI	Late-Life Function and Disability Instrument
LOI	Letter of Intent
MDR	Medical Device Regulation
MDS-UPDRS	Movement Disorder Society-Sponsored Unified Parkinson's Disease Rating Scale
MIF	Maximal Isometric Force
MRI	Magnetic Resonance Imaging
MS	Multiple Sclerosis
MVIC	Maximum Voluntary Isometric Contraction
(N)MSK	(Neuro)Musculoskeletal Model

OFL	Optimal Fibre Length
PCSA	Physiological Cross-Sectional Area
PD	Parkinson's Disease
PFF	Proximal Femur Fracture
PRO	Patient Reported Outcome
QA	Qualification Advice
QO	Qualification Opinion
QT	Qualification Team
RF	Rectus Femoris
RWS	Real World Walking Speed
SaMD	Software as a Medical Device
SAWP	Scientific Advice Working Party
SNMES	Superimposed Neuromuscular Electrical Stimulation
SO	Static Optimization
TSL	Tendon Slack Length
TVS	Technical Validation Study
US	United States
V&V	Verification & Validation
VI	Vastus Intermedius
VL	Vastus Lateralis
VM	Vastus Medialis
VVUQ	Verification Validation Uncertainty Quantification
WB	Walking Bout
WP	Work Package

Table of Contents

Abstract	I
List of the main Abbreviations	III
Table of Contents	V
Thesis overview	1
Chapter 1 – INTRODUCTION	3
1.1. The Human Movement as an interplay of physiological systems	3
1.2. Mobility quantification: previous, current and innovative solutions	5
1.2.1. Wearable sensors	8
1.2.2. Musculoskeletal models.....	10
1.3. The concept of credibility	11
1.3.1. Marketing authorisation of medical products: drugs vs medical devices	12
1.3.2. Software as a Medical Device	17
1.3.2.1. Certification of medical device software with measuring capabilities	18
1.3.2.2. Certification of medical device software with predictive capabilities: Credibility	19
1.3.3. Risk-based credibility assessment: the ASME V&V-40:2018 technical standard	21
1.4. Aims of the thesis	22
1.4.1. Advancing the state of the art: novel contributions of the thesis.....	23
1.4.2. The effect of COVID-19 on the projects	23
Chapter 2 – CREDIBILITY OF MOBILITY MONITORING SOFTWARE	25
2.1. The Mobilise-D project	25
2.1.1. The hardware component.....	27
2.1.2. The analytics software	29
2.1.3. WP5 – regulatory qualification.....	30
2.2. The credibility of software for mobility monitoring in regulatory drug trials	31
2.2.1. Qualification advice procedures with the EMA	32
2.2.2. Qualification program procedures with the FDA	33
2.3. The credibility of software for mobility monitoring for clinical use	33
2.3.1. Mobility monitoring software as a measurement device: metrological validation.....	34
2.3.2. Mobility monitoring software as a predictive device: model credibility.....	36
2.4. Acknowledgements	40
Chapter 3 – FORCELOSS: AN OVERVIEW OF SCIENTIFIC RATIONALE AND ENABLING TECHNOLOGIES	41
3.1. The ForceLoss project	41
3.2. Experimental measurements	46
3.2.1. Magnetic Resonance Imaging.....	47
3.2.2. Maximum Voluntary Isometric Contraction test	47
3.2.3. Electromyography.....	49
3.2.4. Superimposed neuromuscular electrical stimulation	52
3.2.5. Bioelectrical Impedance Analysis	52
3.3. Musculoskeletal dynamics models	53
3.3.1. The idealised features	53
3.3.2. Model personalisation.....	58

3.4. Simulation environment	60
3.4.1. Static Optimisation	60
3.5. Model credibility assessment.....	61
Chapter 4 – FORCELOSS: A SIMULATION FRAMEWORK FOR THE DIFFERENTIAL DIAGNOSIS OF DYNAPENIA BY MEANS OF SUBJECT-SPECIFIC MUSCULOSKELETAL MODEL	63
4.1. Materials and Methods	63
4.1.1. Magnetic Resonance Imaging.....	65
4.1.2. Maximum voluntary isometric contraction test.....	67
4.1.3. Electromyography.....	70
4.1.4. Superimposed neuromuscular electrical stimulation (patients cohort only).....	70
4.1.5. BIA (patients cohort only).....	71
4.1.6. Hand-grip test (patients cohort only).....	71
4.1.7. Modelling pipeline.....	71
4.1.8. Biomechanical simulations.....	73
4.1.9. Simulation environment.....	74
4.2. Results	75
4.3. Discussion	77
Chapter 5 – GENERAL DISCUSSION AND CONCLUSION	80
5.1. On the Mobilise-D project	81
5.2. On the ForceLoss project	82
Appendix A – Description of the Mobilise-D analytics software.....	88
References.....	93
Acknowledgements.....	114

Thesis overview

Hereafter, the structure of this thesis will be outlined, and a broad overview of the chapters will be provided. In brief, over the three years of my PhD, I have been involved in two projects: Mobilise-D and ForceLoss, funded by the European Union and the Italian Ministry of Health, respectively. Despite an apparently different focus, the two projects share the clinical context, i.e., human movement, and the fact that the solutions developed within required a credibility assessment. In fact, both technologies are intended to be used as a tool for clinical decision-making. For the Mobilise-D project, the new methodology consists of the analytics software used to extract parameters (e.g., real-world walking speed) from raw data recorded continuously for seven days by wearable sensors. In the ForceLoss project, the focus is on musculoskeletal models.

The thesis is structured into five chapters as follows.

Chapter 1 – Introduction

A general introduction to human mobility and the interconnection of different physiological human systems is provided. Current and innovative solutions to quantify mobility, giving a brief overview of their history, are then explored. Finally, the concept of credibility (of a new methodology) is introduced.

Chapter 2 – Credibility of mobility monitoring software

The chapter focuses on the Mobilise-D project and the work carried out as part of the regulatory work package (WP5). Following a brief description of the relevant regulatory pathways, both with the European and the United States regulatory agencies, the focus will move to the regulatory science and the activities required and undertaken to assess the credibility of the Mobilise-D analytics software. In particular, two different approaches to demonstrate the credibility of such new technology, which come from different perspectives, will be detailed.

Chapter 3 – Differential diagnosis of dynapenia by means of subject-specific musculoskeletal models

Firstly, the ForceLoss project (context, aims, hypothesis and the experimental protocol) is explained. Then, a general description of all the methodologies (i.e., magnetic resonance images, dynamometry test, neuromusculoskeletal models, etc...) required for the study is provided.

Chapter 4 – Forceloss: a simulation framework for the differential diagnosis of dynapenia by means of subject-specific musculoskeletal model

This chapter provides the methodology of the ForceLoss project. A first validation of the ForceLoss framework, on a cohort of healthy females, is provided, while the application to the patients cohort is just introduced and explored in terms of feasibility and expectations since the enrolment of patients, elected for total knee arthroplasty, is still ongoing.

Chapter 5 – General discussion and conclusion

In this final chapter, a general discussion is provided, followed by closing remarks and a final reflection on the main topic embracing both the Mobilise-D and ForceLoss projects.

Chapter 1 – INTRODUCTION

The first chapter briefly explains the clinical context of this thesis: mobility intended as the capacity to move around and as a result of the intricate interactions between multiple physiological systems. After a general introduction, the focus will be on wearable sensors and musculoskeletal models, that represent the two most innovative solutions to monitor and study human movement. The second part of this chapter revolves around the concept of credibility, meant as the fundamental procedure for gaining trust in a new methodology, concentrating on the credibility assessment of a new medical product and the certification process for the different types of software employed in the medical field. The intention is to provide a basic overview of the regulatory pathways and sufficient details to understand the following chapters. Finally, the aims of the thesis are presented.

1.1. The Human Movement as an interplay of physiological systems

Mobility, i.e., the ability to move oneself around freely and easily [1], is a complex and dynamic interplay of multiple organ systems within the human body. It goes beyond the simple act of walking and includes a wide spectrum of movements essential for activities of daily living (ADLs), functional independence, and overall well-being. As the sixth vital sign [2], mobility represents an important marker of health. Mobility impairment is associated to numerous adverse events [3], and an individual's ability to move is strongly linked to his health status and quality of life [4,5]. Mobility has become even more of a concern now that the world's population is rapidly ageing. In the last decades, the average life expectancy increased of 20 years, and according to the World Health Organization, the number of people older than 60 years is expected to double by 2050 [6].

Mobility is the result of the seamless orchestration of various physiological systems, including the neuromuscular, musculoskeletal, cardiovascular, respiratory, and sensory systems [7]. The basis for human movement lies in the neuromuscular system, which governs the initiation, coordination, and control of movements. The central nervous system (CNS) plays a key role in the control of mobility by exercising its influence through various cognitive, sensory, autonomic, and motor networks [8]. Signals from the brain instruct muscles to contract and relax in a synchronised manner, facilitating precise and coordinated movements. The musculoskeletal system, consisting of bones, muscles, joints, and connective tissues, provides the structural framework for mobility. Bones offer support and protection, muscles generate the force required for movement, while healthy joints enable coordinated motion [9]. The cardiovascular system plays a vital role in mobility by ensuring adequate oxygen delivery to various tissues and organs [10]. The respiratory system, encompassing the lungs and airways, is integral to mobility. Efficient respiratory function manages the exchange of oxygen

and carbon dioxide, ensuring an ample oxygen supply to muscles and thus allowing activities of different intensities [11]. Finally, sensory systems, including vision, proprioception, and vestibular function, contribute crucial information for maintaining balance, coordinating movements, and navigating the surrounding environment [7]. Vision aids in spatial awareness and proprioception provides a sense of body position, and the vestibular system contributes to equilibrium [12]. Due to the intrinsic interplay of these physiological systems, a disruption of one or more systems directly and adversely affects mobility. Dysfunction in the neuromuscular system can impair mobility, leading to conditions such as muscle weakness [13], postural instability [14], or coordination deficits [15]. Moreover, neuromuscular disorders affect the interaction between the nervous system and muscles, impacting the ability to walk, maintain balance, or perform ADLs [13]. A few examples of movement-related disorders due to neuromuscular dysfunction are Parkinson's disease, multiple sclerosis, Huntington's disease, muscular dystrophy and poliomyelitis. Musculoskeletal disorders, like osteoporosis or osteoarthritis, as well as other conditions such as injuries or fractures compromise the structural integrity and function of this system, directly resulting in reduced and impaired mobility [16–18]. Sarcopenia [19], the loss of skeletal muscle mass with advancing age, and the subsequent loss of muscle strength referred to as dynapenia [20] highly affect mobility [21,22]. Cardiovascular diseases, such as congestive heart failure or coronary artery disease, can compromise the heart's ability to supply sufficient blood, adversely impacting mobility. Poor circulation can lead to fatigue, weakness, and impaired mobility [23,24]. Furthermore, decreased endurance, reduced mobility, and difficulties in maintaining physical activities can be caused by high blood pressure or hypertension, as well as by peripheral artery disease and heart rhythm disorders, such as arrhythmias [25]. Respiratory conditions, such as chronic obstructive pulmonary disease or asthma, can limit mobility by causing shortness of breath and reducing overall endurance and balance [26,27]. Impairments in sensory function can lead to difficulties in spatial orientation, affecting the precision and safety of movements [28,29]. Individual's mobility capacity is even more undermined by the combination of systems disruptions. An evident example is the frailty syndrome, one of the most prevalent and significant geriatric syndromes, characterised by age-related declines in physiological functionality, involving multiple organ systems and leading to adverse and unfavourable health outcomes (e.g., higher risk of falls) [30–32]. A reduced and impaired mobility is one of the main negative consequence affecting a frail person, together with weakness, slowness, low level of physical activity, poor endurance and weight loss [31]. Even if the aetiology of frailty is unclear, it has been observed that there might be two possible pathways for the onset of frailty: one as a result of physiological changes associated with ageing that are not related to specific diseases (such as ageing-related sarcopenia); the other as a conclusive common pathway linked to severe disease or comorbidity, as

indicated by elevated rates of poor health status and a more extensive range of subclinical physiological changes observed in frail individuals.

To sum up, as people get older, they are more likely to become frail and to develop chronic disease, musculoskeletal disorders and pathological health conditions (i.e., osteoarthritis, hip fracture, or sarcopenia) that directly affect their ability to move around [1]. Regardless of the underlying health condition, several studies have proven that the loss of mobility or an abnormal gait (e.g., a slow walking speed) are associated with higher fall risk and hospitalisation, greater morbidity and mortality, cognitive decline, and loss of independence [33–38]. Moreover, the consequences of mobility loss, associated with the population ageing, have a great impact not only on the personal life of people, leading to a reduction of the quality of life, but also on the society and the healthcare system. More frequent hospitalisations, higher need assistance and long-term care increase the healthcare costs [3]. Recently, the concept of *healthy ageing* has been introduced and has become of primary importance. Healthy ageing does not mean that people get older free of disease, but it is about creating the environments and the structures to enable well-being and independence in older age. Since the consequences of the loss of mobility, even more so in an older population, are a current and urgent issue, the cooperation of industries, researchers and individuals is required to develop and implement innovative and feasible solutions.

1.2. Mobility quantification: previous, current and innovative solutions

When considering the quantification of mobility, there is often an inclination to focus solely on laboratory tests; however, a comprehensive approach to mobility quantification should encompass various facets, including the perception of movement, to provide a holistic understanding. Recently, three different constructs to describe/characterise mobility have been introduced [39]:

- Mobility capacity. It refers to the ability of an individual to move and the intensity that a person uses while performing an assigned motor task.
- Mobility perception. It captures the patients' subjective experience of their mobility through patient-reported outcome measures or clinical-assigned scores.
- Mobility performance. It refers to the intensity, quality and extension of patients' mobility as observed in the real-world environment and monitored for multiple days or, in general, for a sufficiently long period of time to be considered representative of daily life.

Mobility capacity

The quantification of mobility capacity, aimed at studying and identifying its principles and its biomechanical causes and effects, has always sparked interest. The current gold standard for gait

analysis is the stereophotogrammetric system, which allows precise three-dimensional measurements to be obtained from stereoscopic images. Stereophotogrammetry has roots in the pioneering experiments of Eadweard Muybridge in the late 19th century [40]. Muybridge's setup consisted of 24 cameras placed in line at a constant distance from one to another to produce a detailed and chronological record of the subjects' and animals' motion, posing the basis for photogrammetry methodologies [41]. Muybridge was the first to demonstrate that all four of a horse's hooves are off the ground simultaneously during a gallop (Figure 1.1.A). Another of Muybridge's noteworthy works is *The Woman Walking Downstairs*, which is considered one of the earliest attempts to objectively study the biomechanics of the human body during everyday tasks (Figure 1.1.B).

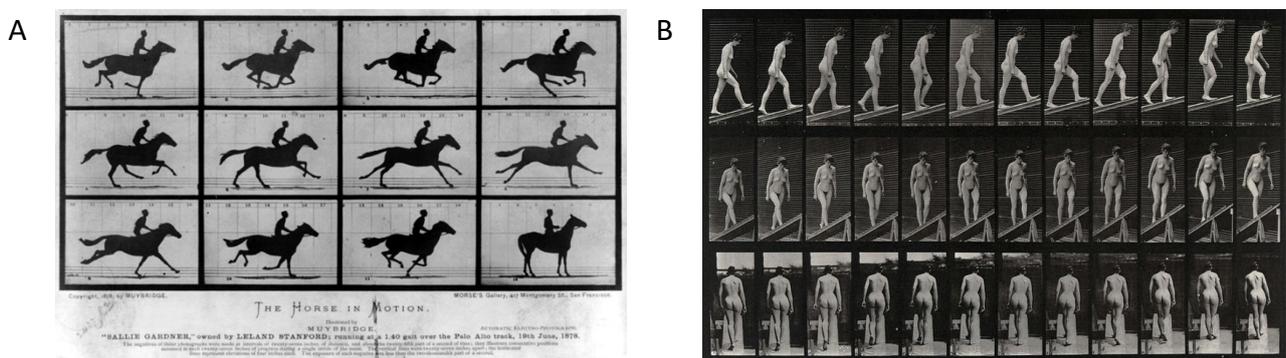


Figure 1.1 The pioneering works of Edwing Muybridge: (A) the Horse in Motion (1878), and (B) the Woman Walking Downstairs (1887).

Few years later, in 1882, Étienne-Jules Marey, a French scientist and photographer, invented the chronophotographic gun, a modified shotgun capable of capturing 12 consecutive frames per second on a photosensitive disc. His most known works focused on animals, such as a study on how falling cats land on their feet (Figure 1.2.A), and secondary on the biomechanics of human walking, further improving existing photographic techniques by incorporating markers on the subject's body (Figure 1.2.B) [42].

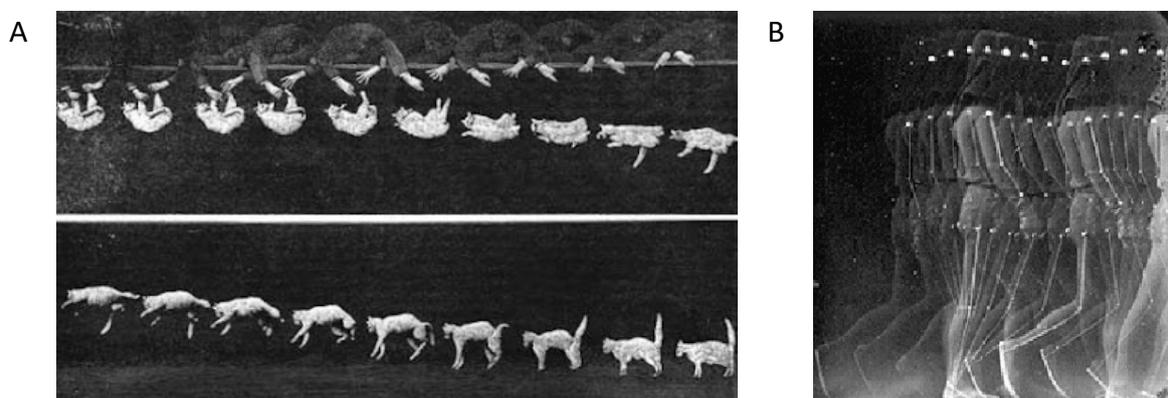


Figure 1.2 The two main work of Marey: (A) the Falling cat (1894), and (B) the Walk (1886).

Finally, the work by Braune and Fischer on human biomechanics and gait analysis, often involving cadaveric dissections and detailed anatomical measurements, laid the groundwork for subsequent advancements in motion capture techniques like the stereophotogrammetry [43]. Nowadays, stereophotogrammetry is the gold standard for motion analysis, allowing for the acquisition of immediate marker positions on the skin surface, utilising either traditional photography or optoelectronic sensors [44]. Stereophotogrammetric systems typically consist of multiple cameras strategically positioned to capture overlapping views of a target area or object (Figure 1.3), allowing for the triangulation of points in space and enabling the creation of three-dimensional reconstructions [45]. Modern stereophotogrammetric systems often incorporate optoelectronic sensors, infrared cameras (e.g., Vicon system [46]), and automated algorithms, contributing to improved accuracy, reduced setup times, and enhanced ease of use [42]. In addition to this technology, for a more comprehensive analysis, dynamometers, like force plates, are used to measure external forces (e.g., ground reaction forces) [47], while electromyography (EMG) allows recording the electrical activity of muscles [48]. In clinical settings, the 6-minute walking test [49], the Timed Up and Go test [50], and the L-test [51] are just a few examples of gait analysis tests used to quantify the mobility capacity of a person. Although these tests allow us to quantify mobility directly, they suffer from inherent limitations.

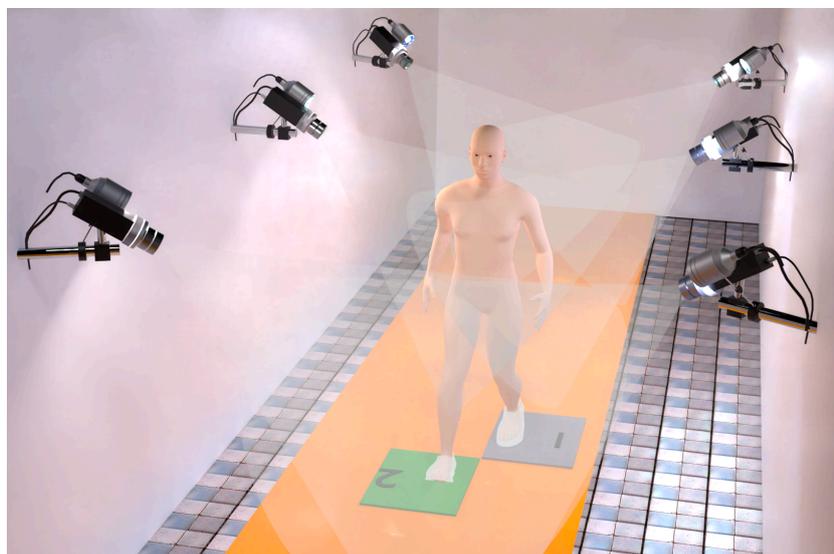


Figure 1.3 Stereophotogrammetric system with two force plates, placed in the centre of the field to acquire the ground reaction forces in addition to the subject's movement.

Typically, clinical tests are conducted in an instructed and supervised environment, lacking the person's genuine intent during the investigation [52]. Moreover, recent studies have demonstrated that these assessments fail to provide insights into the real-life performance of ADLs (e.g., going for a walk, shopping, taking the stairs, or having a bath, and social interactions) [53–55]. Finally, symptoms such as motor and nonmotor fluctuations, early morning dystonia, and various daily life

issues like sleep problems, which are not evident in a clinical setting but significantly impact individuals' experiences, are unlikely to be captured [52].

Mobility perception

In parallel, to quantify mobility perception, various patient-reported outcomes (PROs), which encompass questionnaires and diaries, have been developed. PROs are self-reported measures and thus primarily capture individuals' subjective experiences and perceptions of their mobility and related aspects, like daily function [52]. However, it is important to consider potential biases to which PROs are susceptible, such as the recall bias [56], the testing bias, or the mood influences [57] when interpreting mobility-related PROs. Additionally, cultural background, knowledge, beliefs, and misinterpretations of standardised questions, physical symptoms, and reactions can further influence PROs [58]. Some examples of questionnaires are the Movement Disorder Society-sponsored Unified Parkinson's Disease Rating Scale part II (MDS-UPDRS-II) for Parkinson's Disease patients [59] and the Multiple Sclerosis Walking Scale-12 (MSWS-12) for multiple sclerosis cases [60].

Mobility performance

To overcome the limits of the abovementioned tests, the attempt to find innovative ways for quantifying mobility performance has spurred an intense research activity, which has produced a new generation of wearable sensors, able to quantify mobility continuously for multiple days [39,61]. Moreover, in parallel with the advances in wearable technology, another innovative methodology (i.e., musculoskeletal modelling) for studying human movements has arisen.

1.2.1. Wearable sensors

The definition of wearable technology has varied throughout the years since its development from the early 13th century till nowadays, and researchers approaching the concept from various perspectives contribute to distinct interpretations in this field [62]. The first agreed definition dates back to the mid-20th century, with the first modern wearable computer, a timing device hidden in a shoe capable of accurately predicting the landing place of a roulette ball [63,64]. From the 2000s, wearable technologies, also referred to as wearable devices, are compact electronic mobile devices designed to be worn on the body, often integrated into clothing and accessories (e.g., smart watches, bracelets, rings), or in their mini-invasive versions such as micro-chips or smart tattoos [65] (Figure 1.4). Wearable technologies can be classified according to their application/functionality (e.g., for education and professional sports or in the healthcare system) or based on the wearable device types [65]. The best-known intended use of wearable sensors belongs to the consumer market, as wearable sensors are mainly employed as activity trackers [66,67], for monitoring daily activities like step

count [68], basic heart rate [69], and body temperature [70], with the primary aim of promoting increased physical activity among average users.

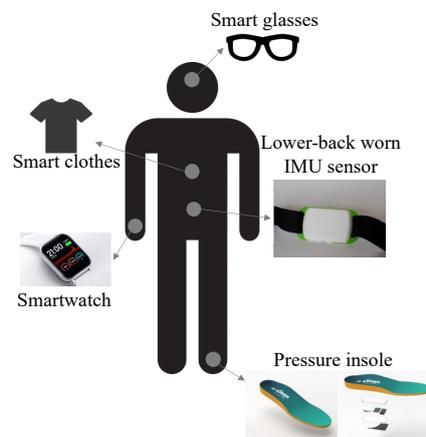


Figure 1.4 Wearable technologies, including smart T-shirt or pressure insoles, highly differ in applications, shapes and body locations.

The great spread of wearable technologies is attributable to their low cost, ease of use, small size, and portability [71]. However, initially, the use of wearable sensors was limited to healthy subjects because of their low reliability and accuracy and the drawbacks that they suffer from [66,72]. These limitations, together with technological advancements (e.g., in miniature devices [73]), have driven intense research activities, which resulted in the development of new wearable sensors that can be used on patients, assisting clinical-decision making. Several clinical studies have been recently conducted to validate the performance of wearable sensors to monitor physiological data over extended durations, aiming to enhance the clinical care of patients (e.g., for patients affected by congestive heart failure [74] or for patients with PD [75]).

Wearable sensors now allow the quantification of real-life mobility for a period of time sufficiently long to be considered representative of daily life [39,61], thus allowing the quantification of mobility performance. Wearable inertial measurement units (IMUs) are the most commonly used sensors for monitoring human movement [76] and consist of accelerometers, gyroscopes and magnetometers (three per type, orthogonally placed, to three-dimensionally quantify movement). This represents the hardware component necessary for recording data. Specific software then elaborates and processes the raw recordings to extract parameters used to characterise mobility, such as average walking speed, cadence, step/stride time, length or turn duration, and angle [77]. To date, there is a lack of standardisation and consensus on the (correct/best) IMU placement, highlighted by the various (body) locations where sensors [78] can be worn on (e.g., ankles [79], wrists [80], or lower back [81]). Among all, the lower back has been recently identified as the ideal sensor location to reliably and consistently measure/extract mobility outcomes as well as the most preferred location by the users [82]. Wearable IMUs used for the gait analysis were initially unable to characterise gait in patients

with slow speed (i.e., gait speed lower than 0.05 m/s) [81,83], since the algorithms for extracting and processing the signals were developed using data from healthy and fit people [72,84,85]. However, a great research activity in this field resulted in the development of the so-called second-generation mobility monitors (i.e., a new generation of multi-sensor IMUs). Although the hardware part of the wearable IMUs remained similar, the software for signal processing has been innovated and become more and more sophisticated to be able to analyse abnormal gait and slow speed [86–88].

1.2.2. Musculoskeletal models

Over the past 20 years, in parallel with significant advances in wearable sensors and miniaturised technologies, another innovative methodology for the study and analysis of human movement has emerged: musculoskeletal (MSK) modelling. MSK models are computational representations of the human body and, specifically, of the MSK system, combining anatomy, physiology, biomechanics, and engineering (see Figure 1.5).



Figure 1.5 Musculoskeletal dynamics model of the lower limb, where each muscle is represented with (one or multiple) actuator(s).

While introducing MSK modelling, it is necessary to make a distinction between generic and personalised (i.e., subject-specific) models. Due to the high complexity of the MSK system, the generation of subject-specific MSK models is far from easy. Initially, only generic MSK models, which were defined using cadaveric data (from a small pool of subjects) [89–96], were used because of the time-consuming process and the amount of data (e.g., medical images, kinematics data) needed to generate personalised models [97]. However, generic MSK models hardly represent a specific individual, let alone children or minorities, as their constitutive properties and characteristics are based on cadaveric data [98–100]. The easier access to subject-specific data (e.g., medical images),

the advancements in the software technologies [101,102] to develop MSK models and in simulation environments [103,104] contributed to a wider (although still limited) adoption of personalised models of the MSK system in various fields and disciplines. MSK models and biomechanical simulations, which fall in the in silico medicine category, have shown to have great potential in orthopaedics, allowing the estimation of quantities which would be impossible to measure without invasive surgery (e.g., joint contact force) [105], and predicting various musculoskeletal outcomes after orthopaedic surgeries [106–108]. Moreover, these models have been employed to explore the impact of sub-optimal control on joint forces [109], and act as a stratification tool in the treatment of juvenile idiopathic arthritis [110,111]. Finally, personalised MSK models have allowed to improve the diagnosis and the treatment of patients affected by MSK disorders [112,113] and used in defining rehabilitation programs [110,114]. However, the wide-scale use of the MSK models is limited by both the complexity of representing the high variability of the human system and, above all, the validation of models and the assessment of their credibility.

1.3. The concept of credibility

Credibility is a multifaceted concept, but its definition varies based on the field where it is applied. In general, the credibility of a new methodology refers to the trustworthiness, and reliability of that methodology, whether it is a new medical product or a MSK model. In the context of new medical products, credibility is central to gaining the trust of healthcare professionals, regulatory bodies, and patients. The credibility of a medical product is intricately tied to its safety, and efficacy. Similarly, in the field of MSK models, credibility is essential for the acceptance and use of these computational representations of the MSK system in the clinical setting. The accuracy of these models in simulating human motion and estimating internal forces, such as muscle and joint forces, establishes their credibility. Rigorous validation against empirical data and continuous refinement contribute to the credibility of MSK models, enabling researchers and practitioners to trust the insights derived from these computational tools.

Historically, the concept of credibility was introduced in the medical field as a foundation upon which trust of a new medical product was built, referring to the extent to which such a new product is considered trustworthy, effective, and safe for its intended medical purposes. A key role in this process is played by health authorities, such as the Food and Drug Administration (FDA) in the United States (US) or the European Medicines Agency (EMA) in Europe, which are responsible for regulatory approval. In general, before a new product can be put on the market, evidence of the safety and efficacy of such a new product for the intended and declared use must be presented to the country's regulatory agency [115]. Evidence has been historically provided through controlled

experiments, referred to as clinical or pre-clinical trials (whether humans are involved or not, respectively). While, traditionally, the assessment of safety and efficacy relied on supervised experiments conducted *in vitro* and *in vivo*, either in animals or in humans, regulatory authorities in the US and in Europe have recently started to consider the use of *in silico* trials (i.e., that rely on computational models and simulations) as a means to provide evidence for regulatory approval [116,117]. However, there is still an ongoing debate regarding how to establish the credibility of *in silico* methods, in contrast with the well-established practices for the credibility assessment of *in vitro* and *in vivo* tests [118–121].

1.3.1. Marketing authorisation of medical products: drugs vs medical devices

Before any new medical product can be marketed, it must obtain the marketing authorisation from regulatory agencies (e.g., EMA and FDA). This authorisation is granted when a company developing and marketing a medical product presents extensive technical, pre-clinical, and clinical evidence demonstrating the safety, efficacy, and quality of the new product. This evidence must be generated using clinical, statistical, or instrumental methodologies deemed appropriate by the regulators [122,123]. The regulatory process to obtain marketing authorisation differs based on the type of medical product. Historically, the main distinction among medical products was between medicines (i.e., drugs) and medical devices. However, the initial clear division between drugs and medical devices started to fade since more complex medical products have been developed (e.g., drug-eluting stents, which are a combination of different types of medical products) [124]. The EMA defines drugs as “any substance or combination of substances presented for treating or preventing disease in human beings or animals. Any substance or combination of substances which may be administered to human beings or animals with a view to making a medical diagnosis or to restoring, correcting or modifying physiological functions in human beings or in animals is likewise considered a medicinal product” [125]. A medical device is instead “any instrument, apparatus, appliance, material or other article, whether used alone or in combination, including the software necessary for its proper application intended by the manufacturer to be used on human beings for the purpose of:

- diagnosis, prevention, monitoring, treatment or alleviation of disease,
- diagnosis, monitoring, treatment, alleviation of or compensation for an injury or handicap,
- investigation, replacement or modification of the anatomy or of a physiological process,
- control of conception” [126].

The choice of the regulatory framework for certifying a new medical product is guided by the country where the new product will be sold, and it differs from country to country. While the principles are

similar, the details can vary considerably (see Table 1.1, where the main steps in the regulatory framework for a medical device, outlining the difference between the American and European pathways, are summarised).

Table 1.1 Differences between the European and US regulatory processes for the qualification and/or marketing authorisation of new medical products.

Regulatory system	CE Marking (Europe) ¹²	FDA approval ³
Risk class of the product	<p>Risk classification is based on 18 rules that guide the identification of which class risk the device is. Moreover, the class of risk of the device is based on the intended use and on the consequences of the clinical decision informed by the software itself:</p> <p>Class IIa: software intended to provide information which is used to make decisions for diagnosis or therapeutic purposes.</p> <p>Class III: if such decisions have an impact that may cause death or an irreversible deterioration of a person's state of health.</p> <p>Class IIb: if such decisions have an impact that may cause a serious deterioration of a person's state of health or a surgical intervention.</p> <p>Class I: All other device.</p>	<p>Risk class is defined mostly by similarity. If a device is "substantially equivalent" to another already FDA-approved, the simplified 510(k) premarket submission can be used.</p> <p>Class I: low-risk devices</p> <p>Class II: medium-risk devices</p> <p>Class III: high-risk devices</p>
Required pre-clinical and clinical tests	<p>The European system is based on a hierarchy of requirements: it starts with essential requirements that all medical devices must have, outlined in Annex I of Directive 93/42; it follows a hierarchical system always based on harmonised ISO standards</p>	<p>Based on the risk classes, there is a precise FDA procedure to follow among the 510(k) premarket submission, De Novo, Humanitarian Device Exemption, and Premarket Approval</p>
Technical Documentation	<p>A technical file with a pre-defined structure must be drafted</p>	<p>Market notification or market approval based on the selected FDA procedure</p>
Implement Quality	<p>According to ISO 13485:2016⁴</p>	<p>FDA requires Current Good Manufacturing Practices (CGMP), with the details of the requirements, and recommends as the best</p>

¹ <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:02017R0745-20230320>

² https://health.ec.europa.eu/system/files/2021-10/mdcg_2021-24_en_0.pdf

³ <https://www.fda.gov/medical-devices/overview-device-regulation/regulatory-controls>

⁴ <https://www.iso.org/standard/59752.html>

Hereafter, a brief mention of the regulatory authorities in Europe and in the US is provided [127]. In the US, the FDA is the unique authority that deals with food, drugs, medical devices, tobacco products, vaccines, cosmetics, and other products. For each product, the FDA has a specific centre to address the requests. For example, the Center for Devices and Radiological Health (CDRH) is responsible for the certification process of medical devices (i.e., by assuring the safety, efficacy, and high quality of medical devices), while the Center for Biologics Evaluation and Research (CBER) regulates biological products for human use. In Europe, regulatory authorities differ on whether the medical product is a drug or a medical device. For drugs with a significant therapeutic, scientific, or technical innovation, the EMA represents the sole regulator for the scientific evaluation, supervision and safety monitoring of medicines in the EU. Otherwise, manufacturer must both apply the European standards and the guidelines developed by national regulatory agencies (e.g., the Italian Medicines Agency in Italy, or the Federal Institute for Drugs and Medical Devices in Germany) for receiving marketing authorisation. For the medical device, the only valid technical standards are produced by the European Standards Organisation, such as the European Committee for Standardization (CEN). The International Organization for Standardization (ISO) represents another important source of technical standards; even if the ISO is not a European standard, the majority of the ISO standards are harmonised and recognised as equivalent to CEN standards.

The qualification process of a novel methodology for drug development provided by the EMA can be pursued through the submission of a request for qualification opinion, often preceded by one or multiple requests for qualification advice. The latter is meant for submitters seeking for preliminary feedback or guidance provided during the early stages of development or submission of a proposal for a novel methodology. A Qualification Advice (QA) is not a formal regulatory decision but serves as valuable input to the developers, helping them to refine their approach before pursuing a formal qualification opinion. A Qualification Opinion (QO) reports the agency's official position on whether a specific novel methodology is considered acceptable for use in the development and evaluation of medicines. This procedure is similar to the Biomarker Qualification Program (BQP), i.e., the US equivalent regulatory pathway for new medicinal products. Hereafter, a more detailed description of the different regulatory qualification procedures in Europe and in the US is provided.

Regulatory qualification in Europe

For the European market, the EMA has devised a voluntary procedure to achieve the regulatory qualification of new (health) biomarkers (e.g., Mobilise-D digital mobility biomarkers or DMOs), as

outlined in the *Qualification of novel methodologies for drug development*⁵ guidance for applicants (and schematized in Figure 1.6).

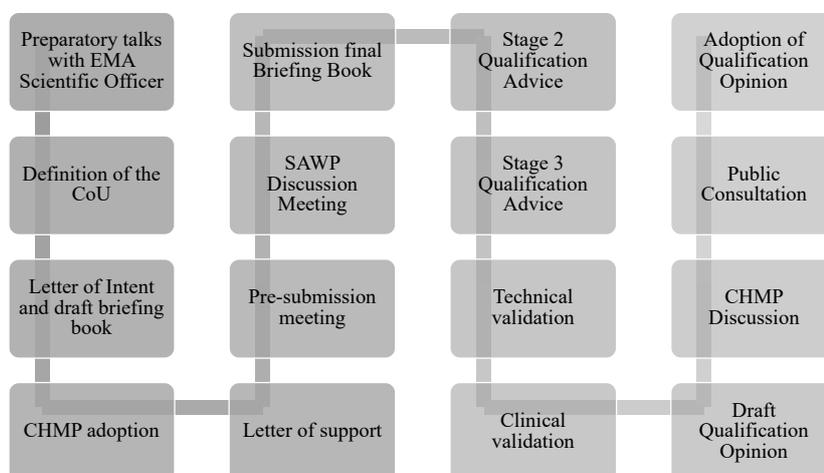


Figure 1.6 The entire workflow for the qualification of a new methodology for medicine development to EMA. Adapted from [14] under the terms of the CC BY 4.0 license. Copyright © 2020 by the authors. (CHMP: Committee for Human Medicinal Products; SAWP: Scientific Advice for Working Party).

The qualification process involves the Scientific Advice Working Party (SAWP) and the Committee for Medicinal Products for Human Use (CHMP) and may result in two potential outcomes: CHMP QA or CHMP QO. The overall duration of the QA process is approximately 160 days, while the QO process takes up to 250 calendar days. The procedure initiates with the submission of a Letter of Intent (LOI) and a draft dossier 60 days prior to the official start of the request, allowing flexibility in deciding between QA or QO pathways. On the regulator’s side, a qualification team (QT), including coordinators (SAWP or CHMP) and subject experts relevant to the technology’s Context of Use (CoU), is typically formed within the first few weeks of application. This marks the official start of the procedure (i.e., day 0), followed by several discussions during QT and SAWP meetings, as well as meetings with the applicant, where a list of questions is issued and discussed, leading to the issuance of a draft report at day 70. On day 90, the SAWP provides a recommendation on whether to pursue the QO or QA pathway based on the request type and available data. If the applicant opts for a QA, this will be adopted by the CHMP at the first CHMP Meeting on Day 100. A Letter of Support may be drafted and published on the EMA website, which summarizes the QA received. Alternatively, if sufficient information is available, a draft QO is issued, discussed, and adopted during the first and second CHMP meetings (day 100 and 130, respectively). Following CHMP adoption, the QO undergoes a 6-week public consultation period to gather opinions from the scientific

⁵ https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-novel-methodologies-drug-development-guidance-applicants_en.pdf

community. By the subsequent 15 days, the QO is adopted and published. The described qualification procedure is not exempted from the fee payment, computed according to the *Explanatory note on general fees payable to the European Medicines Agency as of 1 April 2023*⁶ document. However, discounts apply to Small Medium-Sized Enterprises.

Regulatory Qualification in the US

In the US, the qualification of biomarkers falls under CDER BQP as part of the broader Drug Development Tools (DDTs) Program [123]. DDTs encompass various methods and materials aimed at facilitating drug development, covering biomarkers, clinical outcome assessments, animal models, and more. The BQP process does not involve any fees. Typically, each biomarker undergoes a separate submission for qualification unless multiple biomarkers are intended to be combined to represent a single CoU. The qualification framework consists of three main stages, as depicted in Figure 1.7. An additional initial meeting (stage 0) may be scheduled to interact with the regulatory board before starting the official qualification procedure [123].

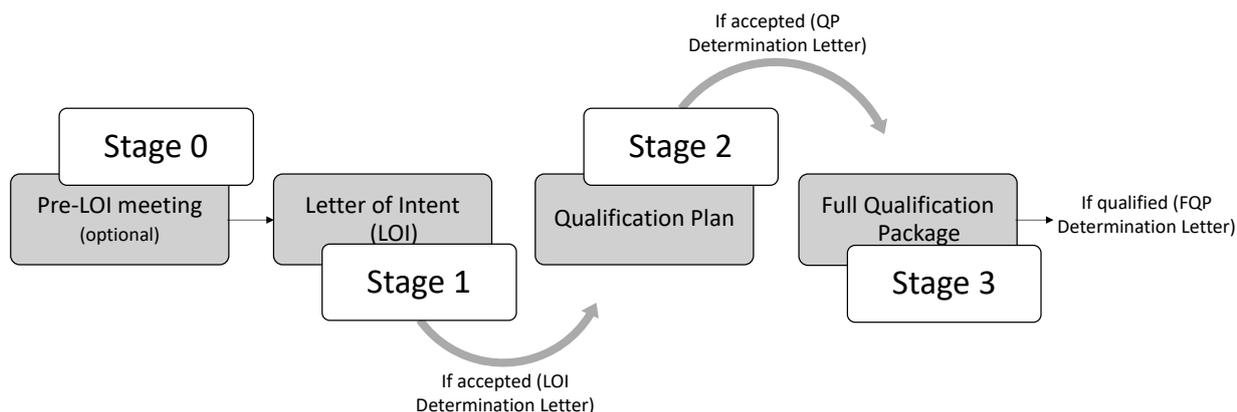


Figure 1.7 The framework of the FDA Biomarker Qualification Program procedure.

The LOI is an initial communication from the applicants, providing an overview of the biomarker, outlining the purpose and goals of the qualification process, and describing the CoU. The BQP, submitted in stage 2, is a comprehensive document detailing the strategy, methods, and study design for validating the biomarker. The BQP provides a roadmap for the qualification process, outlining the scientific and regulatory approach to demonstrate the biomarker’s reliability and relevance for its intended use. The full qualification package includes a detailed report on the validation studies conducted to support the biomarker’s qualification, presenting the comprehensive scientific evidence and data, demonstrating the biomarker’s reliability and its suitability for its intended CoU, based on

⁶ https://www.ema.europa.eu/en/documents/other/explanatory-note-general-fees-payable-european-medicines-agency-1-april-2023_en.pdf

the recommendation from the previous stages. At each stage, a three-step procedure involving initial assessment by the FDA, comprehensive scientific review, and evaluation by the DDT Committee is required to receive the acceptance or the qualification (for stage 1 or 2 and for the stage 3, respectively). The timeframe for the review at each of the three main stages is 3, 6, and 10 months [123].

In the next section, the focus is on medical device software, a recent – yet important – subgroup of medical devices that has been historically excluded from the first definition of medical device.

1.3.2. Software as a Medical Device

According to the International Medical Device Regulators Forum (IMDRF), which is a worldwide group of medical device regulators with the aim of accelerating and harmonising the medical device regulatory pathway, software as a medical device (SaMD) is defined as “software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device:

- SaMD is a medical device and includes in-vitro diagnostic medical device;
- SaMD is capable of running on general-purpose (non-medical purpose) computing platforms;
- “without being part of” means software is not necessary for a hardware medical device to achieve its intended medical purpose;
- Software does not meet the definition of SaMD if its intended purpose is to drive a hardware medical device.
- SaMD may be used in combination (e.g., as a module) with other products including medical devices;
- SaMD may be interfaced with other medical devices, including hardware medical devices and other SaMD software, as well as general purpose software
- Mobile apps that meet the definition above are considered SaMD”⁷.

It is interesting to note that the inclusion of software within the definition of medical devices has evolved over time both in the European and the American context. Hereafter, as an example, the history of the SaMD in Europe is provided. Initially, the software was not considered in the quality assurance process, and the risk factor associated with the software was not taken into account. In Europe, only in the early 1900s, the first standard (i.e., the ISO/IEC 9126 *Software Engineering* –

⁷ <https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf>

*Product Quality*⁸) for the quality assurance of software was published. A few years later, software technologies started to be seen as a continuously varying and updating models, hence a new standard was published: the ISO9000-3 *Quality Management and Quality Assurance Standards*⁹, which introduced the concept of development, supply, installation and maintenance of computer software for assessing the software quality. At the beginning of the 21st century, when the software started to be used to program a medical device (e.g., programmable pacemakers) or represented itself as a medical device (e.g., the surgical planning software in computer-assisted surgery), the medical device definition was expanded by including the software alone (i.e., as stand-alone software) or in combination. Therefore, in 2007, the concept of SaMD was introduced. The Medical Device Regulation (MDR) (EU) defines a SaMD, but referring to medical device software, as the “software that is intended to be used, alone or in combination, for a purpose as specified in the definition of a “medical device” in the medical devices regulation¹⁰ or in vitro diagnostic medical devices regulation¹¹”¹². In this thesis, the terms SaMD and medical device software will be used as synonyms. From a regulatory point of view, the certification framework differs based on the nature of the software (i.e., whether it has measuring or predictive capabilities).

1.3.2.1. Certification of medical device software with measuring capabilities

SaMD with measuring capabilities, e.g., Endosize software (Endosize; Therenva, Rennes, France) [128], is certified following well-established procedures and standards. According to the IMDRF, all SaMD used in clinical settings shall undergo a clinical evaluation (see IMDRF/SaMD WG/N41FINAL:2017 document¹³). The mentioned document represents the basis on which both the European and the American guidelines for the certification process of the SaMD are based (e.g., in the EU context, Medical Device Coordination Group MDCG 2020-1 guidance¹⁴). The clinical evaluation consists of a set of activities for the determination of clinical safety, effectiveness, and performance. In this process, the technical and clinical validation are pivotal. The technical validation quantifies the ability of a SaMD to accurately, reliably and precisely generate the intended technical output from the input data, while the clinical validation seeks to demonstrate that the SaMD can generate clinically relevant output in accordance with its intended purpose. All these procedures are

⁸ <https://www.iso.org/standard/16722.html>

⁹ <https://www.iso.org/standard/26364.html>

¹⁰ Article 2(1) of Regulation (EU) 2017/745 – MDR

¹¹ Article 2(2) of Regulation (EU) 2017/746 – IVDR

¹² https://health.ec.europa.eu/system/files/2020-09/md_mdcg_2019_11_guidance_qualification_classification_software_en_0.pdf

¹³ https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-170921-samd-n41-clinical-evaluation_1.pdf

¹⁴ https://health.ec.europa.eu/system/files/2020-09/md_mdcg_2020_1_guidance_clinic_eva_md_software_en_0.pdf

referred to as metrological methodologies and aim to verify the metrological performances of the device/software. Moreover, because of the changing nature of the software (i.e., it can always be modified or updated), the concept of life cycle of the software was introduced. The CEI/IEC 62304: *Medical device software – Software life cycle processes*¹⁵ standard defines guidelines for the software life cycle processes to ensure the secure development and maintenance of the SaMD. During the life cycle of the software, the manufacturer must:

- keep the clinical evaluation updated, drafting the associated documents, including the Post Market Clinical/Performance Follow-up as described in the MDCG 2020-1 guideline;
- carry out continuous risk management as shown in the ISO 14971¹⁶;
- establish and maintain a quality management system. For this purpose, based on the company and the software product, the two reference standards are the IEC 62304¹⁷ and the ISO 13485¹⁸;
- apply for a type-examination from a Notified Body that shows that the product complies with the provisions of the MDR.

Since the FDA has recognised as consensus standard the international harmonized standard within the EU (i.e., CEI/IEC 62304), all the information described above are also applied to the American context.

1.3.2.2. Certification of medical device software with predictive capabilities: Credibility

The term predictive SaMD encompasses a range of software technologies, from embedded software based on math models with predictive capabilities to digital patient technologies. A digital twin (DT), i.e., a computer model informed with data of a specific subject (subject-specific model) that can predict quantities about that subject which may or may not be difficult to measure directly, represents a relatively recent example of a predictive SaMD. The first certified DT (i.e., the HeartFlow software [129,130]) dates back to 2014. Other examples of SaMD are the in silico trials, predictive software technologies that are not used directly for the treatment of people in healthcare but that contribute to the development (discovery, design) and the risk reduction (including regulatory assessment) of medical products employed for clinical decision-making. In silico methods aim to reduce, refine and replace in vivo and in vitro experiments (1) by limiting the number of animals/patients enrolled in the clinical trial, their duration, and/or the number of measurements conducted during in vitro and in vivo

¹⁵ <https://www.iso.org/standard/38421.html>

¹⁶ <https://www.iso.org/standard/72704.html>

¹⁷ <https://www.iso.org/standard/38421.html>

¹⁸ <https://www.iso.org/standard/59752.html>

experiments, (2) redesigning the study to either eliminate or alleviate the suffering experienced by animals or mitigate the risks for humans participating in experiments, and (3) by replacing the experiments on the animal or human with computational modelling and simulations [131].

Since medical device software with predictive capabilities has started to be used in the healthcare [132,133] to support the clinical decision for a specific subject, the need for a different certification process became apparent and led to the definition of a verification, validation and uncertainty quantification process applied to the medical field.

The general concept of verification, validation, and uncertainty quantification

In many engineering fields (e.g., aerospace, automotive, civil industries) where computer models are used to support product development, well-established procedures, known as Verification, Validation and Uncertainty Quantification (VVUQ, or simply V&V), are used for assessing model credibility [115,121,134]. Hereafter, a brief description of the VVUQ practice and its main components is provided.

- Context of Use. It represents the first step in the process of establishing the credibility of a model. It includes a comprehensive description of how the model will be employed. Precise details must be outlined, including the specific decisions informed by the model, the role of the predicted quantities in these decisions, the structure predictive model, the required inputs and their determination process, the uncertainty associated with these inputs, and the potential range of values assumed by the inputs while using the model in the CoU.
- Verification. Typically, it consists of both code and model verification. Code verification ensures the quality of the software used for solving the model, employing practices from software engineering such as regression tests and the stability of the numerical algorithms. Model verification assesses approximation errors, focusing on quantifying the error resulting from the approximated solution of the mathematical model. Once code verification is completed, any remaining approximation error is attributed solely to the numerical solution – and not to coding errors. Model verification methods vary based on the model’s structure and the numerical methods employed. Usually, for mechanistic approaches, the so-called benchmark tests are conducted to verify the model by solving problems with known solutions, or the model’s results are compared against outcomes computed using already validated software.
- Validation. This is an important step in the process of establishing the credibility of the models. It involves a quantitative comparison between the model predictions and experimental or clinical data computed using a higher-order precision instrument. Validation

becomes even more important when models are used in clinical decision-making and, above all, in high-risk applications since a less accurate model prediction can lead to poorer (or even wrong) clinical decisions.

- Uncertainty Quantification. It estimates how stochastic errors affecting the input propagate through the model into the outputs. Integral to uncertainty quantification, the sensitivity analysis also plays a crucial role in the VVUQ process. It evaluates which elements of the input set are the main drivers of output variability, providing essential insights into the model's behaviour.

Another fundamental element that has a key role in the credibility assessment is the applicability, which refers to the extent to which the model predictions are credible while changing the initial conditions.

1.3.3. Risk-based credibility assessment: the ASME V&V-40:2018 technical standard

The most known technical standards that rely on and develop the VVUQ concept are produced by the American Society of Mechanical Engineering (ASME) standard committee. From the early 2000s, the ASME has defined and published standards to guide the VVUQ processes in different fields, initially in solid and fluid mechanics, and lately (i.e., at the end of 2018) for establishing the credibility of computational models applied to medical devices [135]. In particular, the ASME V&V-40 standards introduce a risk-informed framework for assessing the credibility of a computational model. Unlike other methodologies that serve as guidelines, the emphasis of ASME V&V-40 is on determining “how much” V&V is necessary to ensure model credibility and not on “how to” perform it. The risk associated with using the computational model to inform a decision should determine the level of evidence required. In the ASME V&V-40, the model credibility refers to the trust gained through the collection of evidence in the predictive capability of a computational model [136], and it is established through the verification and validation activities. Verification is defined as “the process of determining that a computational model accurately represents the underlying mathematical model and its solution from the perspective of the intended uses of modelling and simulation”, while validation is “the process of determining the degree to which a model or a simulation is an accurate representation of the real world from the perspective of the intended uses of the model and simulation”. The steps of the ASME V&V-40 framework are briefly described in the following:

- Definition of the question of interest, which is the scientific question to be addressed by modelling.
- Definition of the context of use, the complete description of the planned methodology and the related purpose of use.

- Assessment of the model risk, considering the possibility for the computational model to contribute to decisions that could lead to harm or undesirable outcomes for patients. Model risk is determined by two factors: the model influence (i.e., the impact of the computational model on decision-making) and the decision consequences (i.e., the consequences of adverse outcomes arising from an incorrect decision based on the model).
- Establishing the credibility goals by planning the verification and validation of credibility factors listed in Table 1.2 (some or all based on the risk analysis).
- Perform credibility activities, which include technical and clinical validation.

The ASME V&V-40 is based on the VVUQ methodology, but it introduces variations, such as the risk assessment or the evaluation of human errors (use error, Table 1.2). The additional and final steps involve the draft of the conclusive documents reporting all the evidence for the assessed credibility.

Table 1.2 The 13 credibility factors as presented in the ASME V&V-40 standard represent the goals for assessing the credibility. Reproduced with permission from [136]. Copyright © 2019, Wolters Kluwer Health.

Activities		Credibility factors
Verification	Code	Software quality assurance
		Numerical code verification
	Calculation	Discretization error
		Numerical solver error
		Use error
Validation	Computational model	Model form
		Model inputs
	Comparator	Test samples
		Test conditions
Assessment	Equivalency of input parameters	
	Output comparison	
Applicability		Relevance of the quantities of interest
		Relevance of the validation activities to the context of use

1.4. Aims of the thesis

The aim of my PhD is to understand and, in part, implement the steps required to perform the credibility assessment of two different models: (1) the analytics software used to extract mobility-related parameters from continuous recordings of wearable sensors and (2) the MSK models

employed to simulate a maximum voluntary isometric contraction test of the knee extensor (to identify the primary cause for the loss of muscle force). The work attempts to meet these aims by dividing it into several objectives, such as interacting with the regulatory authorities to define the right certification process and defining a new protocol to perform the differential diagnosis of dynapenia by combining different instrumental measurements.

1.4.1. Advancing the state of the art: novel contributions of the thesis

The projects developed as part of this thesis focused on original and innovative solutions to tackle highly relevant societal themes in healthcare. On one hand, the activities carried out within the Mobilise-D project were aimed to facilitate/promote the use of wearable sensors for continuous movement assessment in the real world to support the development and approval of new drugs to slow down or revert disease progression in clinical indications where (the loss of) mobility is a concern. The focus was on the procedures to certify the device-agnostic multi-module software developed by the technical experts in the Consortium that is capable of extracting mobility-related parameters. Interactions with regulatory bodies have been instrumental in clarifying and defining the steps, previously nebulous and intricate, required for the certification of new medical devices, such as those developed in the Mobilise-D project. The insights gained from these regulatory exchanges have not only streamlined the approval process for the Mobilise-D device but also provided a valuable framework that can aid any entity, including Small Medium-Sized Enterprises, in navigating the complexities of medical device certification. On the other hand, the MSK modelling and simulation framework developed as part of the ForceLoss project by combining experimental data and *in silico* methods promises to enable the differential diagnosis for the loss of muscle force, which is currently not possible – with traditional measures. Peculiar to the project, is the use of MSK models as falsification tool to test different clinical hypotheses.

1.4.2. The effect of COVID-19 on the projects

It is important to emphasise that, as with many other projects, the progression of both the Mobilise-D and the ForceLoss projects was impacted by the unforeseen challenges posed by the COVID-19 pandemic. In the Mobilise-D project, the pandemic and related shortages of electronic components, which limited the availability of wearable sensors, slowed the data acquisition process for the technical validation study, which was crucial for fine-tuning the software pipeline. This challenge necessitated adjustments to the project timeline and methodology, underscoring the need for flexibility in managing research projects during unprecedented global health crises.

Similarly, in the ForceLoss project, besides the data collection challenges, which were already slowed by COVID-19, especially given the involvement of elderly participants, the project faced additional

delays due to equipment delivery issues, further compounded by the global electronics shortage. To mitigate the impact of these delays, for the first part of the ForceLoss study, which saw the enrolment of healthy young adults, an isometric dynamometer – which was readily available – was used instead of the isokinetic dynamometer.

Chapter 2 – CREDIBILITY OF MOBILITY MONITORING SOFTWARE

The current chapter focuses on the credibility assessment of a new medical product, specifically the analytics software developed in the Mobilise-D project to extract mobility-related parameters from real-world activity recordings using a wearable sensor over multiple days. Following a brief overview of the Mobilise-D project and its aim(s), the analytics software pipeline will be described in detail, and two possible regulatory pathways to seek its certification will be illustrated. In fact, while a metrological characterisation is the best-known procedure for assessing the credibility of such a new technology, one could think of applying the ASME V&V-40 technical standards. The latter remains an unexplored pathway, to date, which requires further discussion. A summary of interactions had with the FDA CDRH Division on the feasibility and suitability of such an approach will conclude the chapter.

2.1. The Mobilise-D project

The Mobilise-D project (Connecting digital mobility assessment to clinical outcomes for regulatory and clinical endorsement, www.mobilise-d.eu [39,137]) is a 5-year project funded by the Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 820820 and receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations. The Mobilise-D consortium includes more than 30 international partners among leading international universities, academic hospitals, and some of the world's largest pharmaceutical and technical companies. The foundation of the project is the importance of mobility as an essential marker of health. People's walking ability, specifically when characterised by a slower walking speed, correlates with increased mortality, morbidity, cognitive decline, dementia, and a higher risk of falling [37,38,138,139]. Even though the prevalence of mobility challenges is anticipated to surge with the ageing population, evaluating individuals' mobility, especially in their everyday lives (i.e., mobility performance), is still a challenging task. Mobilise-D is set to establish a comprehensive system for monitoring and assessing people's gait using a single lower-back worn sensor. The Mobilise-D consortium aims to demonstrate that digital mobility outcomes (DMOs) observed in real-world scenarios can effectively predict relevant clinical outcomes, offering an improved, safer, and more expedited approach to advancing the development of innovative medicines. The project concentrates on four different conditions that have in common a direct impact on mobility but different aetiology, such as Parkinson's disease (PD), chronic obstructive pulmonary disease (COPD), multiple sclerosis (MS), and hip fracture recovery/proximal

femur fracture (PFF). The outcomes of Mobilise-D will enhance the precise evaluation of daily life mobility in both clinical trials and patient treatment. This, in turn, will contribute to advancements in care, making it more effective and tailored to individual needs.

Hereafter, a brief description of the patient groups recruited in the Mobilise-D project is provided, together with an overview of the clinically accepted tests to assess mobility capacity and mobility perception for each of the four clinical indications.

Parkinson's disease

Parkinson's disease is a progressive neurodegenerative disorder that predominantly affects movement control. The hallmark symptoms of PD include tremors, bradykinesia (slowness of movement), rigidity, and postural instability [140,141]. As the disease advances, individuals may also experience non-motor symptoms such as cognitive impairment, mood disturbances, and autonomic dysfunction. While the exact cause of PD remains unclear, a combination of genetic and environmental factors is believed to contribute to its onset [142]. Disease severity is typically assessed through the Hoehn and Yahr (H&Y) scale, which is frequently employed to stage the progression of PD based on motor impairment [143]. In clinical studies, the H&Y score is complemented by the revised version of the Unified PD Rating Scale sponsored by the Movement Disorders Society (MDS-UPDRS) [59,144], the current gold standard to assess mobility capacity (Part II) and performance (Part III) in Parkinsonian patients [59]. The latter involves a clinician-assessed evaluation of motor function, providing a standardized method to assess the severity of motor symptoms.

Multiple sclerosis

Multiple Sclerosis is an immune-inflammatory and degenerative disease of the central nervous system, and it is the main cause of non-traumatic disability in young and middle-aged adults [145]. Even if the initial clinical course is highly variable, the onset of mobility problems shows in around 70-95% of the cases [146]. Difficulties in mobility, such as gait impairment, loss of balance and poor coordination, are among the most common deficits in MS patients [147]. Therefore, mobility disability represents the most challenging sign of MS [86] and the major contributor to poor quality of life. In clinical settings, the gold standard for assessing disabilities in MS is the Expanded Disability Status Scale (EDSS), which focuses on ambulatory aspects. The EDSS, typically drawn up by a neurologist, ranges from 0 to 10, with incremental steps of 0.5 and grades 4.0 to 7.5 related to walking disability. Other accepted measures are the timed 25-foot walk component of the MS Function Composite scale and clinical questionnaires (e.g., MS Walking Scale) [86].

Chronic Obstructive Pulmonary Disease

Chronic Obstructive Pulmonary Disease is a progressive respiratory condition characterised by persistent airflow limitation, typically associated with chronic bronchitis and emphysema [148]. One of the key diagnostic measures and indicators of COPD severity is forced expiratory volume in one second (FEV1) [149]. The FEV1 value is a clinically accepted endpoint, useful for assessing the degree of airflow obstruction and for categorizing the severity of the disease (four stages of severity). Individuals with COPD may experience symptoms such as shortness of breath and chronic cough that impact individuals' capacity to move around. Based on the guidelines given by the EMA [150], for COPD patients, cycle ergometry or the 6-Minute Walking Test (6MWT) are the recommended tests to assess mobility capacity. These tests are frequently supplemented by a disease-specific questionnaire (e.g., the COPD Assessment Test [151]) for evaluating patients' mobility perception. An example of PRO widely used to assess the effects of breathlessness on mobility and physical activity is the Saint-George Respiratory Questionnaire [152].

Proximal Femur Fracture

A proximal femur fracture refers to a break in the upper part of the thigh bone close to the hip joint. This type of fracture is common among the elderly population, often resulting from falls or other traumatic events. Proximal femur fractures can significantly impact mobility and independence, requiring prompt and comprehensive assessment for effective management. For PFF, the Short Physical Performance Battery, which consists of a static balance task, a five times chair-raise test, and a 4m walk test at preferred gait speed, represents the clinical test frequently used [153,154]. There is no specific questionnaire to evaluate patients' perception after hip fractures. However, the use of a general quality of life questionnaire (EQ-5D), in which a component is focused on mobility, is accepted [155]. Additionally, it is recommended to assess pain, ADL function, fear of falling and cognitive capacity as co-variates.

Before delving into the description of possible regulatory pathways for the qualification of the Mobilise-D new technologies, in the next sections, the hardware and software components developed by the Mobilise-D consortium will be outlined.

2.1.1. The hardware component

Wearable IMU sensors have been demonstrated to be a novel technology that is low-cost, simple, accurate [156] and capable of being used in the real world, including the home environment and the community. They showed potential for measuring and monitoring real-world walking speed (RWS) and other DMOs [157,158]. However, to ensure reliable and accurate collection of data, the hardware component has to satisfy a set of minimum specifications (e.g., in terms of signal-to-noise ratio, and

sensors' drift). Battery life is of utmost importance, even more so in the context of the Mobilise-D project (aiming at recording data continuously for 7 days). In the Mobilise-D project, the choice of the hardware fell on the DynaPort MoveMonitor (DynaPort MM+, McRoberts, The Hague, The Netherlands. Figure 2.1.A), a class I medical device, CE-marking according to Directive 93/42/CEE, produced by one of the Mobilise-D consortium's partners. The McRoberts device, a lower-back IMU sensor worn at the level of the fifth lumbar vertebra (through a belt), was employed in both the technical validation study (TVS) and the clinical validation study (CVS). The specifications of the McRoberts device (Table 2.1) represented the reference wearable device specifications. The device's wearability and acceptability were taken into account and assessed [159] as patients were asked to wear the sensor for a long period of time and instructed to remove it only during the bath because the sensor was not waterproof [77].

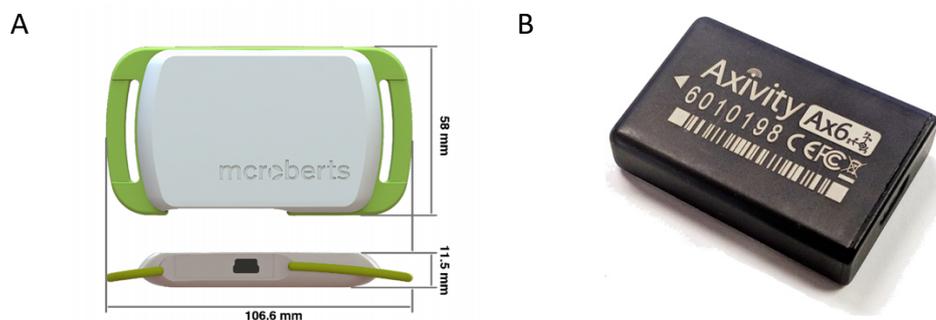


Figure 2.1 Wearable sensors employed in both the TVS and CVS: (A) the McRoberts DynaPort MM+ and (B) the Axivity Ax6.

At a later stage, a second device, the Axivity Ax6 sensors (Axivity Ltd, Newcastle upon Tyne, United Kingdom. Figure 2.1.B), was selected to supplement the stock of Dynaport MM+ to ensure both the TVS and CVS could be completed as planned. The Axivity AX6 sensor was selected as a valid alternative device as it could be configured with the same sampling frequency and sensor ranges reported in Table 2.1, therefore satisfying the minimum metrological specifications required. Moreover, the accelerometers and the gyroscopes included in the adopted wearable device were equivalent in terms of ranges and resolutions to those of the DynaPort MM+.

Table 2.1 Specifications of the McRoberts Dynaport MM+ IMU. Adapted from [77], Copyright © 2021, BMJ Publishing Group Ltd. All rights reserved. Re-use permitted under CC BY-NC.

Dimensions	Weight	Battery life	Sensors	Sampling frequency	Sensor range	Sensor resolution
106.5×58×11.5 mm	55 g	7+ days	Tri-axial accelerometer	100 Hz	±8 g	1 mg (at ±8 g)
			Tri-axial gyroscope	100 Hz	±2000 dps	70 mdps (at ±2000 dps)

2.1.2. The analytics software

The software component allows to extracting the mobility-related parameters (i.e., the DMOs), starting from the raw data acquired using wearable sensors. According to the Mobilise-D approach, the software is independent of the hardware component as long as the latter meets the abovementioned set of specifications (e.g., on the number of sensors, battery life and other metrological characteristics). In the project, the software component consists of the Mobilise-D analytics software, which showed to be highly accurate, compared to gold standard measurements, even when used to monitor subjects who walk slowly or irregularly [77]. This is achieved by transforming what was originally a simple signal processing software into a sophisticated multi-module software, which takes in input the raw signals of the IMU and produces an estimation of various DMOs in output. The Mobilise-D analytics software comprises several modules (Figure 2.2).

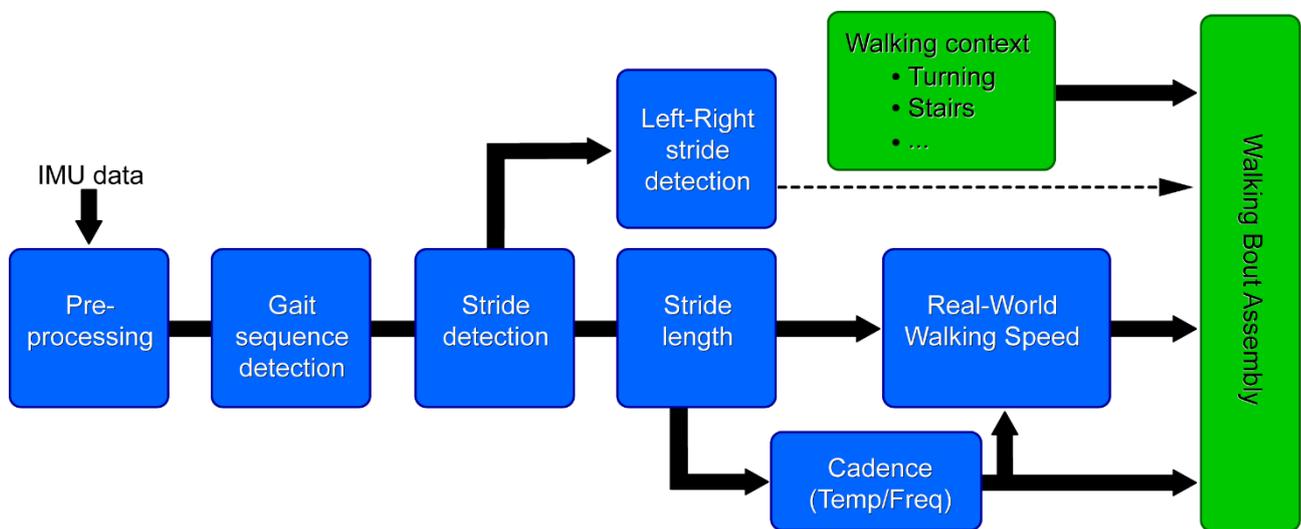


Figure 2.2 The Mobilise-D analytics software for the estimation of DMOs from the raw IMU data.

First, the IMU data are pre-processed to increase the quality of the signals, thereby reducing the noise. This is a critical step, as a wrong upstream analysis would have a negative impact on the efficacy and accuracy of the calculated/extracted DMOs. Once processed, the IMU data enter the gait sequence detection module. As IMUs continuously record for several days at a time, all sorts of activities performed by individuals in their daily lives are recorded. Thus, it is important to isolate the intervals of walking activity for which DMOs will be extracted. Once the gait sequences are isolated, the corresponding data are analysed within the stride detection module. A gait sequence is defined as a succession of strides, each consisting of two subsequent steps, defined in turn by two events (i.e., an initial contact and a final contact between the foot and the ground) [160]. The stride detection module is one of the key modules of the Mobilise-D pipeline, which allows the derivation of numerous mobility-related metrics, such as the stride length, defined as the distance covered by two consecutive placements of the same foot, which is tightly related to cadence and real-world walking speed. If a

finer granularity is required, the data flow in the Right/Left detection module enables the extraction of DMOs at the walking bout (WB) level. A walking bout is a walking sequence containing at least two consecutive strides of both feet (i.e., R-L-R-L-R-L or L-R-L-R-L-R). Moreover, since a WB may be interrupted by resting periods or other activities (e.g., turning, stair climbing), additional algorithms have been included in the Mobilise-D analytics software pipeline (e.g., walking context module, Figure 2.2). For more details, the reader is referred to *Appendix A* at the end of the thesis.

Regulatory strategy: credibility of hardware-independent software

As briefly mentioned in the previous section, while defining a comprehensive approach to assessing mobility based on continuous monitoring in real-world settings, the Mobilise-D project aimed to develop an analytics software that could be independent of the hardware component without losing high accuracy even when used to monitor subjects who walk slowly or irregularly. In this scenario, the software should be certified for clinical use as a SaMD.

2.1.3. WP5 – regulatory qualification

In the Mobilise-D project, seven different work packages (WPs) collaborate to provide validated DMOs, such as RWS, extracted by the Mobilise-D analytics software to monitor disease status and progression in healthy conditions where mobility is a concern from recordings of mobility in real-world scenarios. Among all, *WP5 – regulatory qualification*, is responsible for the regulatory aspects and has the objective of pursuing regulatory qualification for DMOs derived from wearable sensors and encouraging global health authorities and health technology assessment bodies for the regulatory approval in using wearable mobility sensors in the evaluation of safety and efficacy for new medicinal products. Additionally, it seeks to enhance the role of wearable mobility sensors in clinical practice as a technology for both drug development (see Section 2.2) and individual patient care (see Section 2.3). To achieve these goals, WP5 is actively engaging with competent authorities, such as the EMA in Europe and the FDA in the US. In parallel, WP5 investigated the best pathway for the credibility assessment of the Mobilise-D analytics software. Two main regulatory strategies have been foreseen. One is to assess the credibility of the software, considering it with measuring capabilities and thus evaluating it using a metrological methodology. That means the definition of an extensive technical validation protocol [77] followed by clinical validation activities [161]. Alternatively, if the software pipeline is intended as a black box with predictive capabilities, one may decide to assess its credibility according to the ASME VV-40:2018 standards [135]. The effort done so far by WP5 in terms of regulatory qualification of the new methodologies introduced in the Mobilise-D project and the feedback received by the regulatory authorities is described in the next sections.

Personal contribution to the Mobilise-D project

The Mobilise-D project relied on the efforts of a large Consortium comprising both academic and industrial partners all over Europe. As part of the Regulatory Work Package (WP5), I was directly involved in most of the regulatory activities that took place throughout my PhD. While I was only partially involved in the interactions with EMA (i.e., in relation to the two QAs), I had a leading role in the preparation (drafting) and submission of the Pre-submission request to the FDA CDRH division – with the support from the other team members.

2.2. The credibility of software for mobility monitoring in regulatory drug trials

As introduced in section 1.3.1, before any new medical product can be marketed in the European Union (EU), it must obtain marketing authorisation from EMA, as well as from FDA in the American context. To this end, the qualification process of a novel methodology for drug development provided by EMA can be pursued through the submission of a request for qualification opinion, often preceded by one or multiple requests for qualification advice. Instead, the American equivalent process to the QA/QO with the EMA is known as the Biomarker Qualification Program with the FDA. Currently, there are no accepted mobility-related parameters (DMOs) to serve as a biomarker for mobility performance in drug trials [39,162]. The sole exceptions are two recent QOs from the EMA specifically endorsing the use of the DMOs as a secondary endpoint [163,164], highlighting the complexity of the qualification process. To attain such recognition globally, a crucial prerequisite is the demonstration of both technical and clinical validity for each designated CoU within the regulatory framework. The Mobilise-D consortium aims to introduce five distinct CoUs for using DMOs to assess mobility performance across various health conditions. Four contexts are tailored to individual cohorts (PD, COPD, MS, and PFF), while one proposes a disease-independent DMO. Two main different strategies can be pursued to achieve regulatory approval for DMOs, wherein DMOs inform the safety or efficacy of new medical products like drugs or digital therapeutics (as in the Mobilise-D project) [39]. One strategy involves proposing DMOs as primary clinical endpoints to quantify mobility performance directly. Alternatively, a more conservative approach is taken, considering mobility as a secondary DMO, also referred to as a digital mobility biomarker. The latter is relevant for health conditions where mobility outcomes are already accepted for disease monitoring. Demonstrating both technical and clinical validity, including construct validity, predictive capacity, and change detection, is essential for these biomarkers.

Engaging in early dialogues with regulatory authorities, WP5 adopted a staged approach to regulatory qualification, receiving positive advice for a validation protocol, which is further detailed in the following sections.

2.2.1. Qualification advice procedures with the EMA

Based on early interactions with the EMA board, and because of the high complexity of the request, WP5 outlined the qualification strategy, involving a two-part advisory process [162,165].

First request for qualification advice

The first request for QA to EMA was submitted in October 2019, with the main intention of laying the foundation for the general qualification about the use of DMOs as an additional monitoring biomarker in assessing the efficacy of treatments for the cohorts selected in the Mobilise-D project. In this initial stage [162], a preliminary version of the briefing document was submitted together with a letter of intent and to limit the complexity of the request, the defined CoU focused only on PD. The proposed validation protocol was disease-independent, device-agnostic, and based on separating the validation of the device from the validation of the analysis of the data collected using that device [165]. Moreover, the validation included extensive technical and clinical validation involving controlled and unsupervised experiments for the assessment of construct validity, predictive capacity, and the ability to detect change. To this end, two different approaches [39] were implemented: (i) using one disease-specific primary endpoint to reflect disease-related priorities for clinical validation of DMOs and (ii) using a disease-independent outcome, like the Late-Life Functional Disability Instrument (LLFDI), which is an indicator of function and disability in older adults, to inform about the properties of DMOs in different clinical populations with various mobility impairment patterns. Also, secure data management was planned to ensure data integrity in the data transfer, storage, and analysis. A positive qualification advice was obtained [166], in particular, EMA considered acceptable in the proposed validation protocol.

Second request for qualification advice

At a second stage [165], in June 2020, a new request for qualification advice was submitted to the EMA to seek feedback on additional CoUs, where the DMOs were to be used as monitoring biomarkers for the remaining clinical indications of interest (COPD, MS, PFF), to predict clinical outcomes, both global (i.e., LLFDI) and disease-specific. In particular, the disease-specific endpoints proposed by the applicants were the following: fall frequency for both PD and MS patients; occurrence of moderate to severe exacerbations for patients with COPD; and admission to care or nursing home after PFF. The EMA endorsed the general concept of the proposed clinical validation process, which involved 2400 enrolled patients (i.e., 600 subjects for each of the four healthy conditions) in a 24-month observational study. The concept of mobility performance was considered to have the potential to offer valuable supplementary insights into mobility disability in all four targeted diseases: PD, MS, COPD, and PFF. The Letter of Support published in May 2021 [167]

highlighted the lack of an established gold standard to reference, given that mobility performance was a novel outcome without prior evaluation. As a result, the observations might or might not align with other recognised mobility outcomes (i.e., mobility capacity or mobility perception) [167]. Originally, in the light of the advice received, a third submission was planned [39] where the qualification advice on the use of DMOs as biomarkers predictive of relevant clinical endpoints was investigated. In particular, WP5 wanted to seek clarification from EMA experts regarding the best DMO that can be used as a surrogate biomarker to predict (i) for COPD, moderate-to-severe exacerbation necessitating significant changes in treatment or emergency room/hospital admission; (ii) for PFF, prevention of admission to a long-term care facility or other forms of assisted living settings; (iii) for MS and PD: against self-reported falls. However, this third stage has not been actuated yet because it was included in the second request for QA [165].

2.2.2. Qualification program procedures with the FDA

In parallel with the submissions to EMA, WP5 started to approach the FDA board to present the overall project and seek information about the most appropriate regulatory strategy for the Mobilise-D technologies [168]. A pre-LOI meeting has already been held, narrowing the CoU to the patients with MS only. Based on the first feedback received and the two QAs provided by the EMA, the next planned step is to submit a qualification plan to the FDA through the CDER Clinical Outcome Assessment Qualification Program [165], but only once the full clinical studies' results will be available, thus after the end of the Mobilise-D project.

2.3. The credibility of software for mobility monitoring for clinical use

Another aim of the Mobilise-D consortium is to demonstrate that real-world DMOs can effectively predict clinical outcomes, offering a better, safer, and quicker approach for mobility monitoring for clinical use. Currently, there is a lack of robust validation studies for real-world digital mobility measurements, necessitating a technical validation study as the initial step for clinical and healthcare adoption and use. To ensure clinical suitability, the software must obtain certification as a SaMD since it was developed as a standalone component with measuring capabilities, which involves evaluation through metrological methodologies. However, given the close similarity of the software to a predictive model, an alternative approach could involve assessing its reliability using the ASME V&V-40 technical standard [135]. This assessment would be based on the aspects of verification, validation, and uncertainty quantification.

Hereafter, these two possible regulatory pathways are better described.

2.3.1. Mobility monitoring software as a measurement device: metrological validation

Metrological validation refers to the definition of an experimental protocol where well-established tests are implemented to compare the outcomes from the tested device with measurements acquired using gold standard instrumentations or measuring instruments with at least one order of precision higher than that of the tested device. The technical validation protocol proposed by the Mobilise-D consortium (WP2) was defined to compare the DMOs extracted using the Mobilise-D analytics software with measurements acquired using stereophotogrammetry and wearable sensors already validated. In the TVS of the Mobilise-D project [77], a wearable device (i.e., the McRoberts device) and the algorithms have been validated to measure the RWS as a primary DMOs, and other secondary DMOs (e.g., walking bout, stride/step duration, cadence, turning, stride length). The study involved 120 participants from six groups across five clinical sites in three European countries, including patients with COPD, PD, MS, PFF, congestive heart failure, and healthy older adults (HAs). Validation occurred in both laboratory and home settings, assessing the algorithm's viability. The study also included an observational component to evaluate both participants' and professionals' experiences with the device in terms of usability and acceptability. All the gathered information has contributed to defining the inclusion and exclusion criteria for each condition. For all the groups, participants were recruited only if able to walk independently for at least 4 m, read, write, and give informed consent, continuously wear the sensor, and having a Montreal Cognitive Assessment [169] higher than 15, while exclusion criteria (for the clinical cohorts) were whether a medical condition could interfere with the patient's compliance or in the three months prior the inclusion other adverse health conditions had occurred (e.g., myocardial infarction, stroke). See Table 2.2 for the summary of the inclusion/exclusion criteria specific to each cohort; for further details, see *Mazzà et al.* [77].

The TVS involved a comprehensive protocol combining laboratory and unsupervised real-world assessments in three different experimental conditions (see Table 2.3) over nine days: (1) laboratory-based structured activities, (2) 2.5 hours of unsupervised real-world, and (3) continuous monitoring for a period of 7 days. The tested wearable sensor (i.e., the DynaPort MM+) was attached to the lower back using an elastic waistband and Velcro strap. In the in-lab activities, the stereophotogrammetric system was used as a gold standard for the validation of the DMOs during supervised and controlled motor tasks, which mimicked activities of daily life. In the other two contexts of assessment, both in the real-world, the INertial module with DIstance Sensors and Pressure insoles (INDIP) represented the reference system [170–172]. To assess the impact of contextual confounding factors, an additional system, developed as a mobile app (Aeqora App) and capable of detecting outdoor walking, gradient changes during uphill and/or downhill walking, was used.

Moreover, a Bluetooth beacon (BlueBeacon Tag, BlueUp) was used to monitor the activity of the walking aid, when employed [77].

Table 2.2 Summary table with the inclusion and exclusion criteria of the enrolled cohorts. Adapted from [77], Copyright © 2021, BMJ Publishing Group Ltd. All rights reserved. Re-use permitted under CC BY-NC.

Cohort	Inclusion criteria	Exclusion criteria
PD	<ul style="list-style-type: none"> - Aged over 18 years old - Diagnosis of PD based on the Movement Disorders Society criteria 	<ul style="list-style-type: none"> - Impaired mobility related to non-PD causes
MS	<ul style="list-style-type: none"> - Aged over 18 years old - Diagnosis of MS according to the revised McDonald's criteria 	<ul style="list-style-type: none"> - Impaired mobility related to non-MS causes
COPD	<ul style="list-style-type: none"> - Aged over 45 years old - Diagnosis of COPD (post-bronchodilator forced expiratory volume in the first second to forced vital capacity ratio <0.70) - Clinical stability (i.e., at least four weeks without antibiotics and/or oral corticosteroids) - Current or ex-smokers with a smoking history equivalent to at least ten pack years 	<ul style="list-style-type: none"> - With major lung surgery and/or lung tumour - Primary respiratory diseases due to other causes (e.g., asthma) - Impaired mobility related to non-COPD causes
PFF	<ul style="list-style-type: none"> - Aged over 65 years old - Surgical treatment for a low-energy fracture of the proximal femur of the hip and pelvis within the last 12 months 	<ul style="list-style-type: none"> - Impaired mobility related to non-PFF causes
HA	<ul style="list-style-type: none"> - 65+ years of age 	

As a consequent step of the technical validation, extensive clinical validation is also required to have approval by the regulatory authorities. The clinical validation, which is still ongoing, includes the recruitment of 2400 participants (600 for each of the four cohorts selected in the Mobilise-D project).

It is important to note that the TVS represents a huge endeavour, that has to be conducted at every change of the tested device (e.g., an update of the software). These reasons, associated with the high costs require for the TVS, represent non-negligible factors.

Table 2.3 Overview of the experimental protocol employed to validate the algorithms both in laboratory settings and real-world contexts. Adapted from [77], Copyright © 2021, BMJ Publishing Group Ltd. All rights reserved. Re-use permitted under CC BY-NC.

Experimental conditions	Mobility activities	Reference systems
In-lab	<ul style="list-style-type: none"> - Timed Up and Go (TUG) Test (3 meters) - Straight Walking Test - L Test - Surface Test - Hallway Test - Simulated Daily Activities 	Stereophotogrammetric system
Real-world (2.5 hours)	Out-of-lab, free-living in habitual environments chosen by each participant among home, work, and/or outdoor	INDIP Mobile Phone with Aeqora App Beacon
Real-world (7 days)	Unsupervised daily living	Mobile Phone with Aeqora App Beacon

2.3.2. Mobility monitoring software as a predictive device: model credibility

The Mobilise-D analytics software is a highly sophisticated multi-module software that takes in input the raw signals of the IMU and produces an estimation of various DMOs in output. From this perspective, the analytics software could be considered a predictive model that predicts the DMOs starting from raw data acquired by wearable sensors. Therefore, the credibility of the software could be assessed based on the ASME V&V-40:2018 standard [135] (see section 1.3.3).

This represents an innovative regulatory process, and it was thus appropriate to investigate its feasibility and correctness. The Q-submission program [168] developed by the FDA provides an overview of the mechanisms available to submitters through which they can request feedback in writing or during a meeting with the FDA board, and it was identified as the best solution to have interactions with the regulatory authority about the feasibility of such a new regulatory methods. Specifically, a Pre-Submission form, included in the Q-Submission program, which is a voluntary formal written request to obtain FDA feedback prior to an intended submission or to help guide product development and/or application preparation [173], was submitted. The submitters can also decide the method to receive the FDA feedback: written feedback only, teleconference or in-person meeting. The form is exempted from the payment of a fee, and based on the Pre-Submission timeframe reported in the FDA guideline the first written feedback is sent to the submitters by 70 days from the acceptance of the request. A bit longer time (i.e., 90 days) can be required to have also

a face-to-face meeting. In the form, the focus was only on patients affected by MS to reduce the overall complexity of the submission (i.e., narrowing down the Context of Use) and in line with the previous interaction of the Mobilise-D consortium with the US regulatory Agency (see section 2.2.2). The Pre-Submission form consisted of several sections (Table 2.4); the key parts of the dossier were (1) the description of the overall aim of the request, the intended use (of the software pipeline) and the list of specific questions to the Agency (on which we sought to receive feedback), and they are detailed hereafter.

Table 2.4 Summary of the sections and their contents of a Pre-Submission form.

Section	Contents
Cover letter	Contact information Q-Submission type Feedback required (in case of request of meeting, preferred dates and times, planned attendees)
Purpose	The overall purpose of the Q-Sub, including goals for the outcome of the interaction with the FDA
Device or Product Description	Description of device functions and the scientific basis of the device
Proposed Indications for Use or Intended Use	The aim for which the device has been designed
Regulatory History	Listing of any relevant previous communications with the FDA about the subject device
Questions	The core of the form, which includes all the matters the users want to receive feedback about

Aim of the Q-Submission

The aim of the request for feedback for Medical Device Submissions (Q-submission program) was to receive feedback on the regulatory assessment of the Mobilise-D analytics software pipeline when used in the clinical management of patients affected by MS. In particular, the FDA was queried on whether:

- 1) the Mobilise-D software could be submitted as a standalone device and specifically as Software as Medical Device (SaMD) to make it independent from the IMU used to collect the data;
- 2) the appropriate regulatory pathway was a De Novo submission or a Premarket Notification 510(k);

- 3) the ASME V&V-40:2018, intended for the assessment of the credibility of predictive models as medical device development tools, could also be used to evaluate the credibility of predictive SaMD;
- 4) the accuracy of the software in estimating specific digital mobility outcomes should have been formulated as a metrological validation or as the credibility assessment of an orchestration of predictive models. To this purpose, we presented two possible validation pathways for the Mobilise-D software following a metrological validation or following ASME V&V-40:2018 standard.

Proposed Intended Use

The Mobilise-D analytics software is used to process raw signals from patients affected by MS, collected by a wearable Inertial Measurement Unit that satisfies certain technical specifications. The software outputs are the DMOs (e.g., the real-world walking speed) that are used to quantify the mobility performance of the patient, and they will supplement the clinical scales used to evaluate the individual progression of the disease, such as the EDSS. As clinical outcomes, the DMOs computed by the Mobilise-D analytics software can provide additional and supporting information useful for clinical decision-making on the status of MS patients.

Questions

Question #1: Mobilise-D analytics software as a SaMD

The Mobilise-D Analytics Software is designed to be device-agnostic; the software can compute DMOs from the raw recordings of any single wearable IMU as long as it satisfies a set of minimum specifications. Consequently, we believe that the analytic software can be identified as a Software as Medical Device (SaMD) and that can be considered independent from the hardware, as a standalone device. Does the Agency agree?

Applicant's position

In the first two years of the project, the Mobilise-D consortium conducted a rigorous study to show the device-agnostic feature of its methodologies. The TVS, its outcomes and the resulting list of minimal requirements are detailed in peer-reviewed scientific papers [39,77,162]. Leveraging on the results of the TVS, the Mobilise-D consortium would like to treat the analytics software pipeline as a SaMD and plans to assess its credibility independently from that of the hardware component.

Question #2: Certification of the Mobilise-D analytics software

The Mobilise-D analytics software consists of several modules for processing the raw data recorded by a wearable sensor. The intended use is the quantification of the mobility performance in Multiple Sclerosis patients by computing the DMOs, which can be used as supporting tools for clinical decision-making. The premarket notification 510(k) and the De Novo application are two suitable regulatory pathways to certify this software. Based on our research in the 510(k) database, we believe the De Novo request is the right pathway to submit. Does the Agency agree?

Applicant's position

The research in the lists of all the medical devices cleared through the CDRH Premarket Notification 510(k) process led us to believe that De Novo submission could represent the most suitable regulatory pathway to certify the Mobilise-D analytics software, considered as a SaMD. We did not find any previously legally marked device that shares the same intended use as the predicate, and therefore, no substantial equivalence could be demonstrated.

Question #3: Credibility assessment of the Mobilise-D analytics software based on VV-40

The ASME V&V-40:2018 is intended to assess the credibility of predictive models as medical device development tools. Does the Agency agree that the same technical standard can also be used to assess the credibility of predictive software as a medical device?

Applicant's position

The Mobilise-D analytics software pipeline comprises several modules to process the raw data recorded – continuously, over 7 days - by the wearable sensor (hardware component), thus predicting an output (i.e., DMOs). This analytics software consists of both physics-based and data-driven algorithms. As a whole, it could be considered a model that takes in input for the raw data and predicts the DMOs. The ASME V&V40-2018 standard regulates the process to assess the credibility of a predictive model and, as suggested in the FDA guideline [173], could in principle be applied to the several modules of the Mobilise-D software pipeline.

Question #4: Metrological validation vs. model credibility

The Mobilise-D analytics software is a complex pipeline combining signal processing, data-driven modelling, and knowledge-driven modelling. As such, the evaluation of its accuracy and precision in quantifying specific digital mobility outcomes could be conducted in two ways. In the first, we would

assume the software is part of the metrological process, treat it as a black box, and simply compare, in terms of precision and accuracy, for each digital mobility outcome the value provided by the software with that measured by some instrumentation of an accuracy class superior to that we expect for the Mobilise-D software. The second is to treat the software as an orchestration of predictive models and conduct a risk-based verification, validation, and uncertainty quantification analysis for each software module and the whole orchestration separately. Can the Agency advise if at least one of these approaches is considered adequate, and if they are both considered adequate, which one is preferred?

Applicant's position

The most common way to evaluate the validity of the DMOs is the so-called technical validation. It consists of the metrological characterisation of the Mobilise-D analytics software by comparing DMOs predicted by the Mobilise-D pipeline and the same DMOs measured with gold-standard systems, which have an expected accuracy significantly higher than that of the mobility monitors.

A second possibility is to consider the Mobilise-D analytics software as a set of predictive models and, therefore, to assess its validity following the ASME V&V-40:2018 standard.

2.4. Acknowledgements

This study was supported by the Mobilise-D project that has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No. 820820. This JU receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations (EFPIA).

Chapter 3 – FORCELOSS: AN OVERVIEW OF SCIENTIFIC RATIONALE AND ENABLING TECHNOLOGIES

This chapter introduces the ForceLoss project, the second project in which I am actively participating. The primary objective of the project is to establish a novel framework for supporting the differential diagnosis of dynapenia by integrating experimental measurements with computational modelling. My contributions extend to defining the experimental protocol in the laboratory and configuring the acquisition of medical images. However, the significant work I did lies in the computational aspect, including the generation of musculoskeletal (MSK) models and the execution of simulations. Following a brief overview of the clinical context and project objectives, the chapter provides a comprehensive description of all the experimental techniques and methodologies employed in the ForceLoss project.

3.1. The ForceLoss project

The ForceLoss project aims to address this complexity by developing an experimental protocol integrating medical imaging (specifically magnetic resonance imaging – MRI), dynamometry, electromyography (EMG), and subject-specific computer modelling to support the differential diagnosis of dynapenia. With ageing, human muscles naturally become weaker, even in healthy individuals [1]. The loss of muscle force is referred to as *dynapenia* [2], which is both a physiological and/or pathological process commonly observed in the ageing population. Usually, dynapenia is confused with sarcopenia [3], the age-related loss of muscle mass affecting around 10% of elders [4]. This confusion arises from the high complexity of the problem, as the elderly experience functional decline for various reasons, including generalised sarcopenia, fat inclusion due to selective sarcomere loss, activation inhibition (e.g., arthrogenic muscle inhibition after knee surgery), poor neuromuscular control, musculoskeletal disorders or neurological conditions [5,6]. However, the primary causes of dynapenia can be identified in a diffuse or selective sarcopenia, a lack of activation (inhibition), or suboptimal motor control [7,8]. Moreover, muscular weakness causes the elders' physical dysfunction, increased fall risk, and mortality [9–11], thus making early diagnosis as well as the identification of the aetiology and potential treatments crucial. Existing methods, such as maximal isometric contraction test [2], MRI and high-resolution MRI [12], Multi-Channel surface Electromyography (MC-EMG) [13–15], and tetanic Superimposed Neuro-Muscular Electrical Stimulation (SNMES) [13,16], each offer valuable insights into aspects of dynapenia but fall short of

providing an integrated diagnosis. For example, the most common approach to globally examine whether neural deficits contribute to a loss in strength involves the adoption of a SNMES to muscle while the individual is engaged in a maximal voluntary contraction [3].

Commonly, the differential diagnosis of dynapenia begins with a detailed medical history and physical examination, involving a comprehensive assessment to differentiate the dynapenia from other conditions with similar symptoms. Clinicians evaluate the patient's history of muscle weakness, functional limitations, and any associated symptoms such as fatigue or difficulty with activities of daily living. To gather valuable information, diagnostic tests may be performed. These include muscle strength testing using dynamometry, assessment of muscle mass through techniques like dual-energy X-ray absorptiometry or bioelectrical impedance analysis (BIA) [23,24], and evaluation of muscle function and performance using standardised tests such as the Timed Up and Go test [25] or the Short Physical Performance Battery [26]. Laboratory investigations may be conducted to rule out underlying medical conditions contributing to muscle weakness, such as inflammatory myopathies, neurological disorders, or endocrine abnormalities [27,28]. Additionally, imaging studies such as MRI [29] or EMG [14] may be utilised to assess muscle structure and function. To this end, in cases of dynapenia, the necessity to pinpoint the primary cause is crucial due to the variety of medical specialists involved depending on the underlying issue. For patients suspected of sarcopenia, a referral to a geriatrician or an endocrinologist may be appropriate. These specialists are adept at managing age-related physiological changes and hormonal imbalances. If neuromotor deficits are suspected as the cause, a neurologist would be the specialist of choice, given his expertise in disorders that affect the nervous system and its control of muscle functions. A clinical neurophysiologist would be consulted for issues specifically related to muscle innervation. Each of these specialists brings a distinct set of skills and knowledge, making it imperative to accurately diagnose the root cause of dynapenia to ensure the most effective treatment plan.

The ForceLoss project, which aims to conduct a differential diagnosis of dynapenia, combines established diagnostic tests (e.g., dynamometry, EMG, BIA, and MRI) used in such analyses with innovative techniques previously unexplored, such as neuromusculoskeletal (NMSK) dynamics models. The use of subject-specific computer models, often referred to as digital patients or digital twins emerges as a promising solution for complex differential diagnoses. Once validated, these predictive models can simulate various scenarios, aiding in differential diagnoses with high explanatory power due to their mechanistic nature. While the development of subject-specific NMSK dynamics models dates back several years, recent applications have showcased their effectiveness in clinical decision support. These models have been employed to explore the impact of sub-optimal control on joint forces [17], act as a stratification tool in the treatment of Juvenile Idiopathic Arthritis

[18,19], and predict various musculoskeletal outcomes after orthopaedic surgeries [20–22]. The integration of these models into clinical practice represents a significant step towards enhancing diagnostic accuracy and guiding effective treatments in cases of dynapenia.

Two different clinical trials were specifically designed for the ForceLoss project and consisted of the enrolment of 20 healthy adults (*Personalised Modeling and Simulation Procedures for the Differential Diagnosis of Dynapenia: a Study on Healthy Volunteers*. Clinical Trials ID: NCT05091502¹) and 20 total knee replacement patients (*Personalised Modeling and Simulations for the Differential Diagnosis of Dynapenia: Study on Patients With Osteoarthritis (ForceLoss II)*. Clinical Trials ID: NCT05795348²), respectively. All the participants enrolled in the study signed informed consent prior to participating in the study, in accordance with the local Ethical Committee approval (CE AVEC: 216/2020/Sper/IOR).

The general common aims of the ForceLoss project are twofold:

1. develop, optimise, and assess an experimental protocol for performing dynamometry tests aimed at estimating the maximal voluntary isometric force of the muscle groups involved in knee flexion-extension, as well as the relative activation of various muscle groups through EMG recordings and its difference from the maximal tetanic isometric force induced with SNMES;
2. develop, optimise, and (eventually) validate a modelling and simulation framework for generating subject-specific NMSK dynamics models, informed by the data from a MRI of the lower limbs and EMG signals recorded during the dynamometry test, employed for the simulation of such dynamometry test.

Table 3.1 summarises with greater detail the differences in terms of hypothesis and aims of the two ForceLoss clinical studies, referred to as *ForceLoss: Part I – Healthy Volunteers* and *ForceLoss: Part II – Osteoarthritic Patients*, respectively.

Table 3.1 Overview of the two ForceLoss clinical trials, focused on healthy volunteers and osteoarthritic patients, respectively.

	ForceLoss: Part I Healthy Volunteers ³	ForceLoss: Part II Osteoarthritic Patients ⁴
Ages	Adults: 20-40 years old	Older adults: 60-85 years old
Hypothesis	The use of subject-specific NMSK models to simulate a MVIC test of the knee extensors, informed by experimental measures may be employed in the differential diagnosis of dynapenia	

¹ <https://clinicaltrials.gov/study/NCT05091502>

² <https://clinicaltrials.gov/study/NCT05795348>

Aims	<ul style="list-style-type: none"> - Experimental protocol; - Modelling and simulation framework, assessing both feasibility and reliability of the proposed procedures. 	<ul style="list-style-type: none"> - Experimental protocol previously developed and tested on healthy volunteers with additional measures (involuntary muscle contraction, a hand-grip test, bio-impedance and clinical questionnaires); - Modelling and simulation framework to include one additional step (to check for muscle inhibition).
Inclusion criteria	Body mass index: [15-30 kg/m ²]	<ul style="list-style-type: none"> - Diagnosis of Primary Arthrosis at the knee (according to the American College of Rheumatology criteria), subjects elected for total knee arthroplasty; - Body mass index: [18.5-30 kg/m²]; - Health status: 1 or 2 (according to the American Society of Anesthesiology classification); - Suspected systemic sarcopenia due to aging or localized sarcopenia due to disuse.
Exclusion criteria	<ul style="list-style-type: none"> - Neurological, rheumatic or tumoral diseases; - Pathologies or physical conditions incompatible with the use of magnetic resonance imaging and electrostimulation (i.e., active and passive implanted biomedical devices, epilepsy, severe venous insufficiency in the lower limbs, pregnancy); - Previous interventions or traumas to the joints of the lower limb. 	<ul style="list-style-type: none"> - Neurological, rheumatic or tumoral diseases, diabetes, inguinal or abdominal hernia; - Severe Hypertension (Level 3) and/or cardio-pulmonary insufficiency; - Diagnosis of Osteonecrosis in the lower limb joints; - Pathologies or physical conditions incompatible with the use of MRI and electrostimulation (i.e., active and passive implanted biomedical devices, epilepsy, severe venous insufficiency in the lower limbs); - Previous interventions or traumas to the joints of the lower limb.

³ <https://clinicaltrials.gov/study/NCT05091502>

⁴ <https://clinicaltrials.gov/study/NCT05795348>

Each subject has to attend an experimental session to measure the maximum voluntary isometric contraction (MVIC) of the quadriceps and, for the patients, also the tetanic isometric contraction enforced by the electrical stimulation. For each individual, a subject-specific MSK model is generated, with an increasing level of personalisation of muscle properties (i.e., the maximal isometric force – MIF – and the maximum activation level). The proposed framework for the digital twin-assisted differential diagnosis of dynapenia includes the following steps (see Figure 3.1):

- If the experimental MVIC torque is significantly lower than that of the age-matched healthy population, the subject is affected by dynapenia.
- If, for the patients affected by dynapenia, even with the inclusion of the patient-specific muscle maximal isometric force values (computed using the patient-specific muscle volumes segmented from the medical images of the individual), the digital twin predictions are higher than those experimentally measured, then the dynapenia is entirely due to sarcopenia.

- If, even with the inclusion of patient-specific muscle activation levels (calculated from the EMG measurements), the predicted torque remains higher than the measured torque, it is plausible to conclude that the subject, in addition to some degree of sarcopenia, is also affected by some neuromuscular problems (e.g., muscle inhibition or suboptimal motor control).
- If the predicted torques are comparable to that obtained by forced SNMES contraction, then the dynapenia is entirely due to activation inhibition. Otherwise, if, even with the inclusion of the electrical stimulation, the MVIC torque predicted by the model overestimates the experimentally measured one, then the dynapenia is due to problems in the motor control. To confirm this, an EMG-assisted simulation will be run while defining the activation of each muscle based on the measured EMG signals and including a patient-specific suboptimal control. That should bring the predicted value close to the measured torque and confirm that the patient also has neuromuscular control problems.

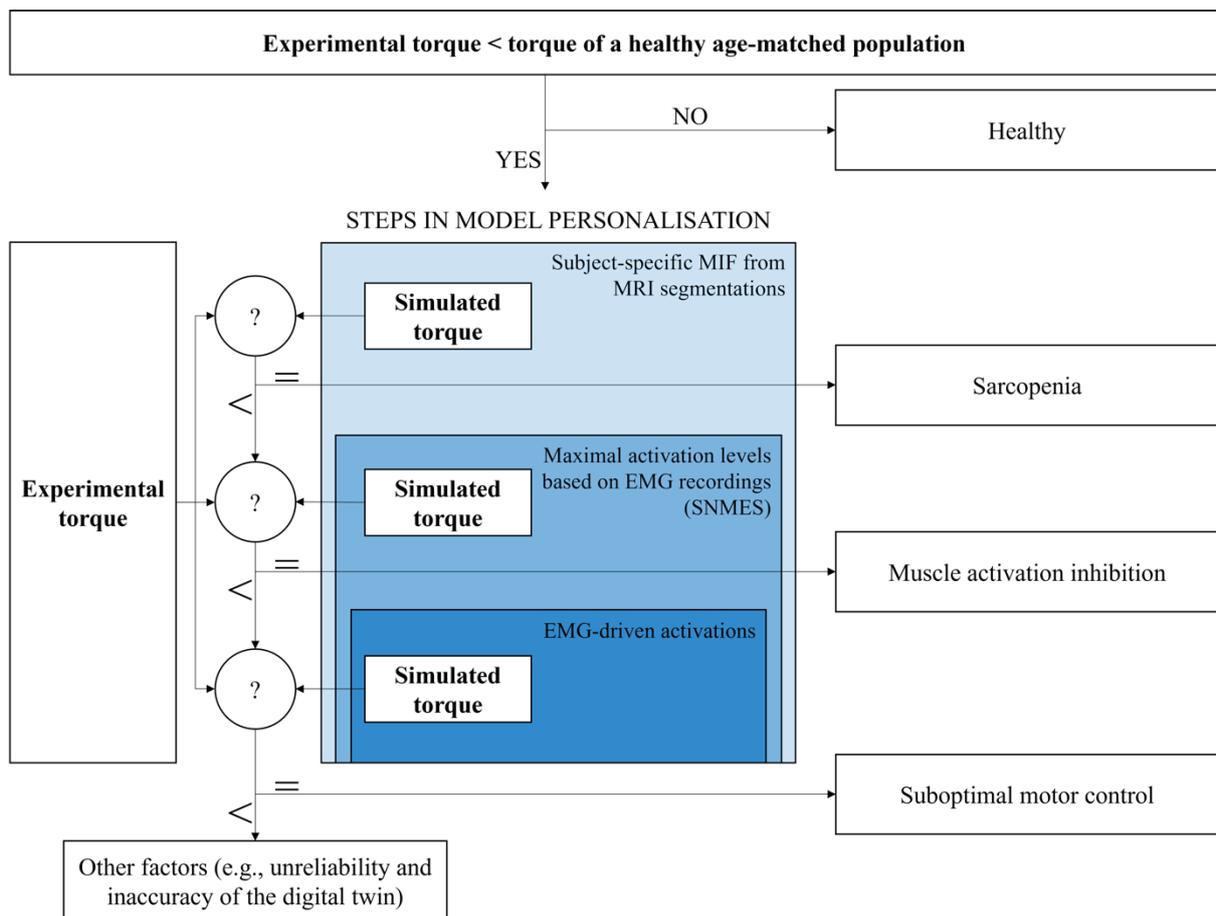


Figure 3.1. The framework for the differential diagnosis of dynapenia, proposed in the ForceLoss project

Personal contribution to the ForceLoss project

My primary responsibilities within the project included data processing, model development, and the execution of simulations. Additionally, support was provided to the research team in data collection, ranging from MRI to laboratory testing.

3.2. Experimental measurements

The ForceLoss protocol (mentioned above) includes the acquisition of an MRI of the lower limbs and the performance of an MVIC test of the knee extensors and flexors while the EMG signals are recorded. Differences in the designed protocols for the two clinical trials are summarised in Figure 3.2. A doctor first visited both the volunteers and the patients to verify their eligibility for the study. The visit also included familiarisation with the instrumentations (in particular, the isometric chair dynamometer or the Biodex dynamometer for the healthy adults and the patients, respectively). After the enrolment, each subject had to do an MRI of the lower limbs, which took approximately 30 minutes, and, on a separate day within the following week, to perform the dynamometry test (after placing the electrodes for surface EMG and the warm-up exercises). The dynamometry test consisted of the repetition of three contractions of the knee extensors and flexors at four (for the volunteers) or at three (for the patients) different knee flexion angles. In addition to the warm-up exercises, upon request of the clinical partners, patients were administered two clinical questionnaires (i.e., Western Ontario and McMaster Universities – WOMAC – and Karolinska Sleepiness Scale – KSS) to collect information on the level of pain and functionality of the joint perceived by the patients, and were asked to perform a hand-grip test as well as to undergo a bioelectrical impedance analysis.

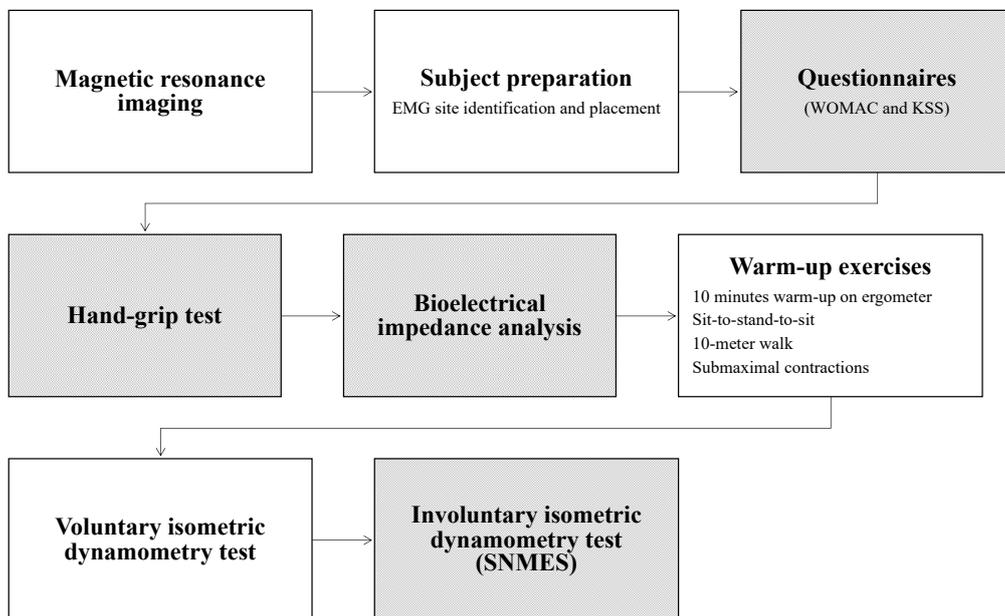


Figure 3.2 The ForceLoss framework designed for the two clinical studies (in the white boxes the clinical tests in common between ForceLoss Part I and Part II are shown, while the grey boxes include the examinations performed only by the osteoarthritic patients).

The following sections briefly describe the various clinical and experimental tests and measures performed in the ForceLoss clinical studies.

3.2.1. Magnetic Resonance Imaging

MRI is a non-invasive, high-resolution medical imaging technique that is considered the gold standard for in vivo studies of muscular morphology and body composition, together with Computed Tomography (CT) [30]. Unlike other imaging modalities such as X-rays or CT scans, MRI does not involve ionising radiation. Despite these advantages, MRI has drawbacks, such as the high cost of scanners [31]. Furthermore, individuals with metallic implants cannot undergo MRI due to safety concerns and potential image distortion (i.e., artifacts) [32]. This restriction extends to those with claustrophobia, given the confined space inside an MRI scanner [33]. Additionally, the quality of MRI images relies on the subject's ability to remain motionless for extended periods (typically 20 to 90 minutes). Even minimal movements, such as tremors or shifts, can compromise image interpretation [34]. Various MRI acquisition sequences have been developed over the years to emphasise specific details based on clinical interest and application. In the ForceLoss project, the proprietary Dixon sequence was chosen for its high versatility and because of the different subjects enrolled in the study, using a GE Healthcare 3.0T scanner.

Dixon sequence

The Dixon sequence offers a versatile approach to imaging by providing four distinct series of images. This sequence initiates with the acquisition of in-phase and out-of-phase images, which are subsequently combined to reconstruct images exclusively highlighting either fat or water. The Dixon sequence's adaptability is particularly advantageous, enabling its application for various purposes, such as the segmentation of different tissues. One notable benefit is its ability to capture microscopic fat infiltration. This sequence has gained increasing interest due to its comparative advantages over traditional imaging technique.

3.2.2. Maximum Voluntary Isometric Contraction test

The MVIC test is a standardised method for the measurement of muscle strength during a static contraction. This test is particularly useful in evaluating an individual's maximal strength within a specific muscle group (e.g., knee flexors or extensors, hip ab/adductors), and it is commonly employed in the sports science [35], rehabilitation [36], and exercise physiology to determine baseline strength, track progress, or identify muscular imbalances [37]. Accurate and consistent MVIC testing provides valuable information for designing personalised training programs and rehabilitation protocols, quantifying overall muscular performance, or monitoring the force loss in patients affected by neuromuscular disease. During the MVIC test, participants are instructed to exert their maximum muscle contraction while maintaining an isometric position (i.e., a fixed joint angle and moment arm). Several methods and instrumentations can be employed for measuring muscle strength. The choice

of the technique depends on factors such as specific research or clinical goals, available resources, and the environment in which the testing will be conducted. Independently from the instrumentations, the MVIC test can be affected by inaccuracies due to intrinsic (e.g., muscle pain [38], personal ambition [39]) and extrinsic (e.g., test protocols, verbal encouragement, visual feedback) factors [40]. Successful testing relies on the patient's ability to fully cooperate and exert maximum effort. Conditions during testing, like fatigue, may also affect the reliability of results, and motivating patients to maintain effort throughout the entire session can be challenging.

Isokinetic dynamometer

Isokinetic dynamometers (IKDs), such as the Biodex (Biodex Medical Systems, Shirley, New York) dynamometer, are sophisticated devices designed to measure and assess muscle strength and joint performance, providing precise control over the speed of movement, ensuring a constant angular velocity throughout the range of motion [41]. The IKDs are equipped with computerised systems that precisely regulate the speed of movement, allowing for concentric, eccentric and isometric contractions, providing a comprehensive understanding of muscle function and potential weaknesses [42]. Moreover, they enable several configurations for assessing muscle function in different joints, including the knee, shoulder, and ankle. Their high precision, reproducibility and reliability, as well as their versatility of isokinetic dynamometers, make them the gold standard for measuring the MVIC test [43–45]. These devices are particularly valuable in clinical and research settings, providing detailed information about muscle strength imbalances, joint stability, and functional capacity. Isokinetic testing is commonly employed in rehabilitation programs [41] for various musculoskeletal conditions, as it allows clinicians to assess and monitor progress while customising treatment plans based on accurate and reliable data. The equipment allows for the isolation of specific muscle groups (e.g., knee flexors or extensors). Straps and belts prevent compensatory movements in other body segments and preserve the alignment between the dynamometer and the rotational axis of the joint of interest, ensuring reliable data acquisition [46,47]. However, wide-scale use is limited due to high cost, large dimensions, and the lack of portability. These can be limiting factors, especially for smaller clinics or facilities with budget constraints or in certain settings, such as home-based evaluations. Isokinetic dynamometers also require regular maintenance to ensure accurate and reliable results. This includes calibrations, software updates, and periodic checks of mechanical components, adding to the overall operational costs [48].

Isometric dynamometer

The isometric chair dynamometer (like the COR1, OT Bioelettronica, Turin, Italy) is a specialised equipment designed for assessing and measuring isometric strength in various muscle groups,

particularly those involved in knee extension/flexion [44,49,50]. The device typically consists of an ergonomic chair equipped with a load cell, adjustable straps to secure the user's body, and a digital display or computer interface to provide real-time feedback on force exertion [210]. Isokinetic dynamometers remain the gold standard even for isometric measurements, but isometric chair dynamometers represent a cheaper but more reliable alternative [216].

In the ForceLoss project, both the Biodex and the COR1 dynamometers have been employed for the MVIC test of the knee extensors/flexors. The hand-grip dynamometer has been used by the patients only.

Hand-grip dynamometer

The hand-grip dynamometer is a portable device designed for assessing the strength of the hand and forearm muscles [180,217]. Consisting of a handle equipped with an adjustable tension gauge, this tool measures the force applied during a maximal isometric grip. The JAMAR Hydraulic Hand Dynamometer (Model J00105, Lafayette Instrument Company, United States of America) is the most frequently used dynamometer and has been considered the gold standard against which other devices are validated [218,219]. One of the primary advantages of the hand-grip dynamometer is its ease of use and portability. Additionally, hand-grip dynamometers are relatively cost-effective compared to more complex muscle strength assessment tools [220]. However, the simplicity of this device is accompanied by limitations. The results may be influenced by factors such as hand position and the individual's technique, introducing variability [218]. Furthermore, hand-grip dynamometers predominantly focus on specific muscle groups, providing valuable insights into upper body strength (i.e., strength assessments of hand and forearm strength) but not a comprehensive evaluation of overall body strength.

3.2.3. Electromyography

EMG is an experimental technique that evaluates and records the electrical activity produced by skeletal muscles in response to nerve stimulation, standing at the intersection of physiology and technology [221]. This diagnostic technique offers a profound insight into the electrical activity of muscles, allowing measurement of muscular performance [222,223], helping in decision-making both before and after surgery [224] and in understanding neuromuscular disorders [225,226], guiding rehabilitation strategies [227], and improving personalised sports activities [228]. The principles of EMG are grounded in the intricate relationship between nerves and muscles, involving the detection and amplification of the electrical signals produced by muscle fibres in response to neural stimuli. As a muscle contracts, motor units are recruited, and the motor neurons emit electrical impulses known as action potentials [229]. EMG electrodes placed on the skin or inserted into muscles capture these

electrical signals, which are then amplified and displayed as waveforms. The amplitude, frequency, and duration of these signals provide valuable information about the muscle's health, activity, and potential abnormalities. Based on the type of EMG, which can be surface or intramuscular [230], there are different types of electrodes:

- skin surface electrodes (for surface EMG): these are non-invasive electrodes placed on the skin above the muscle of interest and embedded in adhesive pads. Their main limitation is that only superficial muscles can be detected. Signal quality can also be affected by external factors such as skin impedance, sweat, and movement artefacts, leading to potential interference and inaccurate recordings. Moreover, surface electrodes may pick up signals from neighbouring muscles, resulting in cross-talk artefacts, where the electrical activity of one muscle interferes with the recording of another;
- needle and fine wire electrodes (for intra-muscular EMG): these electrodes are inserted directly into the muscle tissue, offering more precise recordings by capturing signals from deeper muscle layers and allowing the recording of a single muscle activity. While more invasive, the intramuscular electrodes are affected by little cross-talk concern. However, the repeatability of the measure by repositioning the needle or the wire is nearly impossible. Finally, fine wire electrodes offer enhanced mobility and are often employed for dynamic studies involving joint movements;

Typically, bipolar electrode configurations and differential amplification are used for kinesiological EMG measures.

EMG processing

The EMG processing consists of several steps, including rectifying and differently filtering, which is required to extract the EMG envelope from the raw data. Prior to amplification, the amplitude range of the EMG signal is 0-10 mV (+5 to -5) [231]. The main steps of the EMG signal processing include a pre-filtering for eliminating the various types of noises that contaminate the raw signal (e.g., the dominant concern for the ambient noise arises from the 60 Hz (or 50 Hz) radiation from power sources) [232], followed by rectification, extraction of the envelope and the (optional) normalisation. Hereafter a more detailed explanation is explored.

Raw EMG signals are initially filtered, usually using a zero-lag 4th order Butterworth bandpass filter with 10-500 Hz cut-off frequencies [233]. The frequency lower than 10 Hz are usually cut, since the frequency of the motion noise ranges usually between 1 to 10 Hz and has a voltage comparable to the amplitude of the EMG [232]. Subsequently, the full wave rectification represents a commonly used pre-processing procedure to convert to positive amplitude, transforming all negative amplitudes into

positive values by relocating negative spikes to the positive domain or reflecting them across the baseline. This not only enhances readability but also facilitates the application of standard amplitude parameters, such as mean, peak/max value, and area, to the curve, given that the raw EMG typically has a mean value of zero. The next step is the smoothing of the signal, and two main approaches can be implemented: the linear envelope and the root mean square. For the former [234,235], a zero-lag 4th order Butterworth lowpass filter with a cut-off frequency around 4-6 Hz is used [236], while the latter reflects the power of the signal and has a window with a time duration that ranges between 20 ms (in case of fast movements like jump, reflex studies) and 500 ms (for slow or static activities) [228]. The last step in the signal processing is the amplitude normalisation. This is an optional step implemented to overcome one of the primary limitations in the EMG analysis: the strong influence of detection conditions on microvolt-scaled amplitude data, leading to significant variability between electrode sites, subjects, and even repeated measures. The outcome of normalisation methods leads, therefore, to the elimination of detection condition influence, resulting in rescaled data presented as a percentage of the selected reference value. Different normalisation approaches exist, including normalisation to the internal mean value, a specific trial, or the EMG level of a submaximal reference activity. The most popular one is the maximum voluntary contraction (MVC) normalisation, where the EMG signal is finally divided by the maximum EMG amplitude recorded during a participant's MVC of the muscle of interest. For example, according to [228], the MVC test for knee extensors is performed by performing a leg extension while the subject is seated with the knee flexed between 70° and 90°. MVC normalisation helps account for individual variations in muscle strength and electrode placement. Moreover, it allows researchers and clinicians to compare and analyse EMG signals across different subjects, sessions, or experimental setups, providing a relative measure of muscle activation. In the ForceLoss project, the EMG signals were processed as shown in Figure 3.3 following the steps described above. As the task of interest (MVIC test) was performed in isometric conditions, as suggested in the literature [285], the RMS envelope was deemed more representative (than the linear envelope) and ultimately used to inform the subject-specific models. The envelopes were then normalised - for each muscle - to the maximal value observed across different tasks (i.e., sit-to-stand and MVIC tasks).

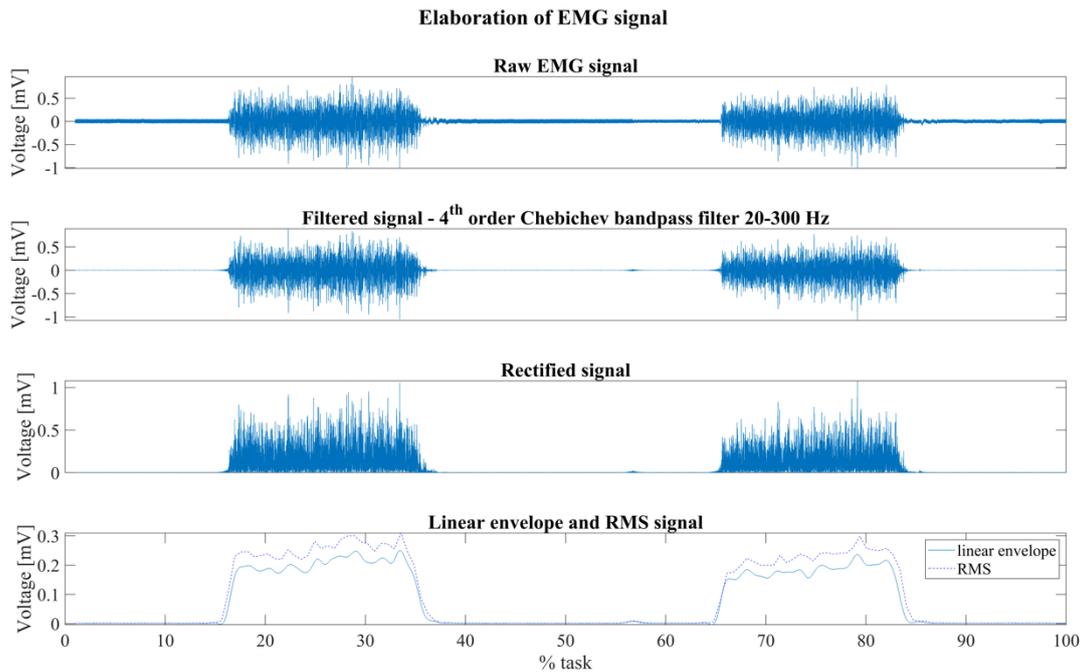


Figure 3.3. EMG signal elaboration from the raw data to the RMS or linear envelope.

3.2.4. Superimposed neuromuscular electrical stimulation

SNMES is an experimental method which involves applying electrical impulses either to nerves or muscles concurrently engaged in voluntary muscle contraction, effectively overlaying external stimuli onto the existing neural drive [237]. SNMES is typically applied using two electrodes positioned on the skin over a muscle belly or muscle group, specifically targeting the motor points [238]. The targeted application of SNMES allows for a refined and individualised approach to rehabilitation [239,240]. Moreover, this technique is adapted to a wide range of therapeutic applications by incorporating it into exercise training programs to optimise strength gains [237] and improve muscle function [241]. However, even if the SNMES has shown a strong potential in enhancing as well as preserving muscle function, there is still a lack of clear consensus on the effectiveness of such a technique [210,215,216]. The stimulation is not without limitations, including patient discomfort, muscle fatigue, and muscle damage [238,244].

3.2.5. Bioelectrical Impedance Analysis

BIA is a method used to assess body composition by measuring the impedance of electrical flow through body tissues. The basic principle is that different tissues (such as muscle, fat, and bone) have different electrical conductivities, allowing for the estimation of body fat and lean body mass [245]. During BIA, a small, safe electrical current is passed through the between surface electrodes typically placed on hand and foot [245]. The impedance encountered by the electrical current as it travels through the body is then used to calculate various body composition parameters, such as resistance,

reactance, total body water, fat-free mass, and body fat percentage. BIA is a quick, non-invasive, inexpensive and relatively simple method for assessing body composition, commonly used in clinical settings and research studies [246]. However, it's important to note that BIA measurements can be influenced by factors like hydration status, body position, and variations in electrical conductivity among individuals [247]. Multiple studies have demonstrated the relationship between dynapenia and body fat mass [190,248]. In particular, individuals diagnosed with dynapenia present with a higher body fat mass in the legs compared to their healthy counterparts. This underscores the significance of BIA as a clinical test to assess the presence of dynapenia.

3.3. Musculoskeletal dynamics models

The MSK dynamics models are rigid segments connected by idealised joints and actuated by muscles. Human models are mathematical representations of the human body and its structures (e.g., skeleton and musculature) that allow the analysis of human movement and quantities (e.g., muscle forces and joint reaction forces) that can be difficult (or even impossible) to measure experimentally. Like any model (e.g., fluid dynamics modelling, mechanical models), the MSK dynamics model is an approximation of the human body, and it thus implies heavy simplifications. Hereafter, I list the most important idealisations and their consequences and their advantages.

In modelling, while introducing simplifications, users must be aware and always keep in mind the maximum level of idealisation that does not compromise the prediction accuracy of their own model. While defining a model, there is continuous research on the balance between idealisations and computational costs. Increasing the complexity of a model a priori does not always represent the best solution, but modelling with a higher level of detail depends on the model application.

3.3.1. The idealised features

Modelling human movement uses the mathematical approach referred to as rigid multibody dynamics, which represents the human body as a set of infinitely rigid bodies linked by joints.

The bodies

The bodies can represent a specific bone or a group of bones (e.g., in an MSK dynamics model, the body tibia usually consists of two bones, the tibia and the fibula). Bones are idealised as infinitely-rigid segments, each characterised by physical dimensions, mass, centre of mass position and moment of inertia. The inertial properties of each body take into account both the skeletal structure (i.e., the bones) and the corresponding body segment, which includes soft tissues like muscles, skin and fat. The infinitely-rigid assumption implies that each segment does not deform under any loading condition. From a kinematic point of view, this simplification allows to characterise the motion of

each body by the description of the motion of a single reference system. Each unconstrained body segment has, in theory, six degrees of freedom (DoFs) in the three-dimensional Euclidean space (i.e., three translations and three rotations). If modelling the bones more accurately (i.e., considering bone deformation when loaded), the computational cost of a dynamic simulation (e.g., gait analysis) would be too high to make the model impossible to use. However, depending on the application, this assumption can be valid or seen as an oversimplification that might lead to inaccurate results. In addition to simplifying human movement analysis, bone idealisation also simplifies the model of joint mechanics by limiting joint movements to single translations/rotations or a combination of them. Based on the data available, body dimensions and properties can be obtained through two main approaches: by scaling them from a generic model (constructed from anthropometric cadaveric datasets) using subject-specific anthropometric measurements or by segmenting them from medical images (e.g., computed tomography – CT – or magnetic resonance imaging – MRI).

The joints

The joints link consecutive bones and represent human internal articulations (e.g., elbow, wrist, hip, knee, ankle). In human bodies, joints are complex structures made up of bones, muscles, and other soft tissues and have multi-axial movements. In MSK modelling, joints are non-deformable and frictionless, and, as a result, forces and moments are transferred through the joints in equal and opposite ways. Moreover, the idealised joints present zero and infinite stiffness, respectively, in the unconstrained and constrained directions, thus reducing the DoFs of each body segment. The majority of the lower-limb joints are synovial, whose mechanical behaviour is viscoelastic (i.e., joint stiffness depends on the velocity) and whose kinetic friction coefficient is around 0.003 [249] and as low as 0.002 [250]. The model's prediction is, therefore, unlikely to be significantly affected by the assumption of frictionless articulation. Moreover, the resultant frictional forces are of lower-order magnitude and thus have negligible respect to the other forces (both internal and external) involved in human movement. In the MSK systems, the joint locations and orientations are typically defined based on well-established guidelines, an example of which is the one provided by the International Society of Biomechanics (ISB). In general, the location of each joint is then defined relatively to the two rigid segments it links and to the DoFs it allows. The joint location and orientation are defined based on the anatomical definitions following the ISB recommendation [251]. The most important joints of the lower limb are idealised, as shown in Table 3.2.

If the ball-and-socket joint for the hip is a common feature of all the lower-limb models, the idealisation of the knee as a planar joint with a prescribed translation between the femur and tibia

based on the flexion-extension angle is not shared by the different models. Some models also include the intra-extra rotation and the ab-adduction in addition to the flexion/extension movement [252].

Table 3.2 The most commonly implemented idealisation of the most important joints of the lower limbs.

Lower-limb human articulation	Idealised joint	DoFs
Hip	Ball-and-socket	3 DoFs: intra-extra rotation, ab-adduction, flexion/extension
Knee	Planar	1 DoF: flexion/extension
Ankle	Hinge	1 DoF: flexion/extension
Subtalar	Hinge	1 DoF: flexion/extension

The muscles

Finally, bones are moved by muscles. The complexity of muscles, in terms of their anatomy, their way of acting, and their functions, requires the assumption of several great idealisations. Firstly, muscles connect just bones to one another. Secondly, the three-dimensional aspect is always represented by one-dimensional unilateral actuators. This assumption can be considered valid for fusiform muscles, whereas several actuators have to be defined for representing a single pennated muscle that thus acts along different lines of action (e.g., in MSK dynamics models, the gluteus medius is usually represented by three actuators). Thirdly, muscle contractile function and force generation are characterised just by mechanical properties, while the complex interaction between these structures and the nervous system is completely excluded. In MSK modelling, the minimal information required for the definition of a muscle is the geometry and the maximal force that a muscle can generate. The muscle path is defined by two main points, the origin and the insertion, respectively. Additional points (i.e., the via points) and parametric surfaces (i.e., the wrapping surfaces around which the muscle line of action wraps) can be added to make the muscle path more physiological and prevent it from penetrating the bones. Other features that characterise each muscle are pennation angle, optimal fibre length, tendon slack length, and activation dynamics (e.g., activation and deactivation times). Over the years, muscles have been modelled differently (e.g., Huxley-type model [253], continuum model [254], non-linear analytic model [255], Hill-type model [256]). However, in the majority of MSK dynamics models, the skeletal muscles are represented as a Hill-type muscle model (e.g., Thelen [257] or Millard [258]). Since this model is computationally less expensive than its alternatives, it is the model implemented by the widely used MSK modelling software (e.g., OpenSim [103,259] and AnyBody [104,260]).

Since the subject-specific MSK dynamics models that I have generated implement the *Millard2012 Muscle Model*, which is based on the Hill-type muscle, a description of this muscle model is provided in the next paragraph. However, for a more detailed and comprehensive description, refer to the systematic review by Caillet et al. [256].

The Millard's muscle model

In Millard's models, musculotendon actuators are modelled assuming they are massless, frictionless and extensible strings, with all the fibres aligned, parallel, of equal length and coplanar. Moreover, to mimic the biological property of muscles to keep their volumes constant, when the muscle is stretching, its area and height remain constant while the pennation angle (i.e., the angle the fibre makes with the tendon, α) varies (Figure 3.4).

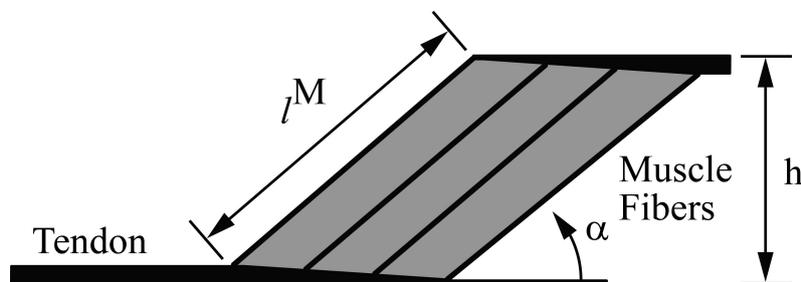


Figure 3.4 Schematic geometric representation of muscle fibres and tendons employed for musculotendon modelling. Adapted from [258].

As mentioned above, Millard's models are based on the generalised Hill-type muscle model, which consists of three main components (Figure 3.5): an active contractile element (CE) in parallel to a passive elastic element (PE) and another passive spring that represents the tendon (T). The CE represents the muscle fibre responsible for contraction and relaxation movements. Also referred to as an active element, it generates the active force of muscles and is thus related to the muscle length. The passive PE takes into account the elasticity and the activation-independent force generated while the muscle is stretched. Finally, the T element represents a passive component and accounts for the activity of tendons and other elastic tissues responsible for the quick length change.

The behaviour of Hill-type based muscle models is described by four characteristic curves: active-force-length, passive-force-length, force-velocity and tendon-force-length curves (Figure 3.6). In MSK modelling, muscles exhibit a behaviour highly dependent on the fibre length and the fibre velocity. The active force of a muscle (Figure 3.6.A) varies non-linearly with the length of the muscle itself; it has its peak (f_o^M) at the optimal fibre length (l_o^M) and decreases while the muscle varies its length (both lengthening and shortening). Differently, as shown in Figure 3.6.C, the passive force exponentially increases when the muscle length exceeds its optimal fibre length, producing force

regardless of whether the muscle is activated. The force generated by muscle fibres also depends on the shortening velocity of the fibres themselves.

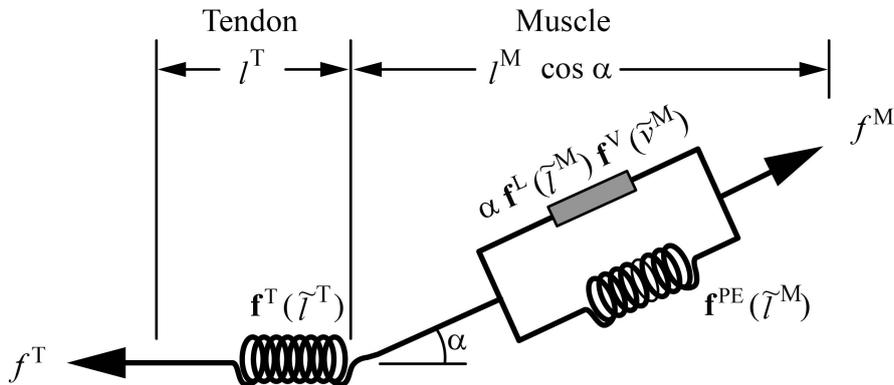


Figure 3.5 The Hill-type musculotendon model components include f^T (tendon force), f^M (muscle force), l^T (tendon length), $l^M \cos \alpha$ (length of the muscle in the direction of the tendon, with α as the pennation angle), $f^T(l^T)$ (tendon-force-length curve), $f^{PE}(\tilde{l}^M)$ (passive force-length curve), and the contractile element (CE) – $\alpha f^L(\tilde{l}^M) f^V(\tilde{v}^M)$ (active force-length curve, where α is the muscle activation, $f^L(\tilde{l}^M)$ is the active force-length curve, and $f^V(\tilde{v}^M)$ is the force-velocity curve). The figure is adapted from Millard et al. 2013 [251].

With a relation represented in the force-velocity curve ($f^V(\tilde{v}^M)$) shown in Figure 3.6.B, during non-isometric contractions, the muscle generates a force which varies non-linearly with its rate of lengthening (the tilde is used to indicate forces, velocities, muscle lengths, and tendon lengths that are normalised by f_o^M , v_{max}^M , l_o^M , and l_s^T , respectively). Muscles and bones are attached through tendons, which are modelled as non-linear elastic elements and generate the force according to the tendon-force-length curve (Figure 3.6.D).

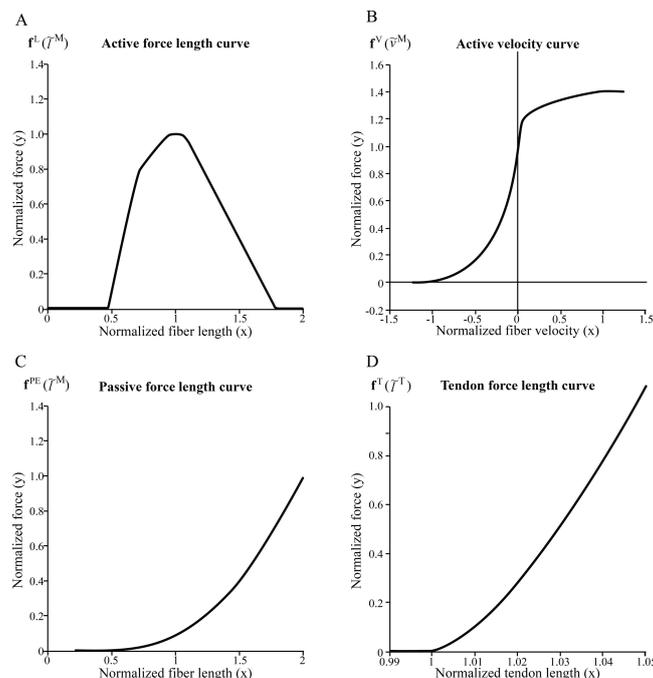


Figure 3.6 Musculoskeletal curves characteristic of every Hill-type based model: (A) active force length curve, (B) force velocity curve, (C) passive force length curve, and (D) tendon force length curve. Adapted from [251].

The model described above is referred to as the *Equilibrium Musculotendon Model*, in contrast with the other Millar's models (i.e., the *Damped Equilibrium Musculotendon Model* and the *Rigid-Tendon Musculotendon Model*). For the exact equations describing these models or any further information, refer to [258].

3.3.2. Model personalisation

MSK dynamics models can be categorised as *generic models* or *personalised models* based on the level of detail they incorporate concerning different aspects such as bone geometries, mass properties, joint models, muscular pathways and properties, and neuromotor control models implemented in the neuromusculoskeletal models.

Generic model vs personalised

Generic MSK dynamics models are computational models that integrate physiological parameters developed from cadaveric databases and are thus representative of a generic population [89–93]. Until the last decade, generic MSK models were more commonly employed, as model personalisation relies on accurately modelling skeletal anatomy, muscle architecture, and decoding neuromuscular control, which is not trivial and requires specialized skillset and dedicated tools. To make generic models more representative of a specific individual, various scaling methods been suggested to incorporate physiological fidelity, also integrating in vivo observations. The main advantage of the generic models is the straightforward implementation compared to subject-specific models, as they do not require individualised data such as subject-specific anatomy, muscle properties, or joint kinematics (even if some measurements like motion capture of standing trial are required). This ease of implementation made them particularly useful for large-scale studies or simulations. However, a significant limitation is their lack of individualisation since they do not account for inter-individual variability, limiting the accuracy of predictions for specific individuals [261,262]. As a result, pathological conditions or certain populations, like children, are not adequately captured by generic models [263]. This limits their effectiveness in studying conditions where individualised anatomical details play a crucial role.

In recent years, personalised MSK dynamics models have gained increased prominence, largely attributable to the growing availability of advanced tools and technologies. The surge in computational power, coupled with advancements in imaging techniques and software tools for generating MSK dynamics modelling, has facilitated the creation of more sophisticated and individualised MSK dynamics models. These personalised models offer a higher degree of accuracy by tailoring the representation of an individual's anatomy and biomechanics, allowing for a more precise simulation of their musculoskeletal system. Additionally, the expanding accessibility of

medical imaging data, such as MRI and CT scans, has provided researchers and clinicians with the necessary inputs for generating personalised MSK dynamics models. The use of these advanced tools not only enhances the accuracy of the models but also enables a more comprehensive understanding of individual variations in musculoskeletal structure and function, paving the way for personalised diagnostics, treatment planning, and rehabilitation strategies. The model personalisation can be done on different levels. Concerning the skeletal model, bone segmentation allows for more personalised joint definitions. For the muscular properties (e.g., MIF, musculotendon lengths), which highly influence the model's behaviour, several methods exist and have been implemented (see Table 3.3 for the different implementation of the MIF), highlighting the complexity of the problem and the lack of standardisation. For example, if muscle segmentations are available, the MIF is usually defined according to the following equation:

$$MIF_m = \sigma \frac{V_m \cos \alpha}{OFL} \quad \text{Eqn. 3.4}$$

where V_m is the volume of the muscle m , α is the pennation angle and σ is the specific tension.

Table 3.3 Various implementation of the maximum isometric force (MIF) proposed by the different authors.

MIF definition	Description	Reference
$MIF_{gen} \frac{m}{m_{gen}} \frac{l_{gen}^{MT}}{l^{MT}}$	<p>MIF_{gen} = MIF of the generic model</p> <p>m = subject's mass</p> <p>m_{gen} = mass of the generic model</p> <p>l_{gen}^{MT} = musculotendon length of the generic model</p> <p>l^{MT} = musculotendon length</p>	Correa et Pandya, 2011 [99]
$MIF_{gen} \left(\frac{H}{H_{gen}} \right)^2$	<p>MIF_{gen} = MIF of the generic model</p> <p>H = subject's body height</p> <p>H_{gen} = body height of the generic model</p>	Steele et al., 2012 [264]
$\frac{V_M}{OFL} \sigma$	<p>V_M = muscle volume</p> <p>OFL = optimal fibre length</p> <p>σ = specific tension $\left(60 \frac{N}{cm^2} \right)$</p>	Handsfield et al., 2014 [265]
$\frac{(V_M V_T)}{OFL} \sigma$	<p>V_M = muscle volume fraction</p> <p>V_T = total leg volume</p> <p>OFL = optimal fibre length</p> <p>σ = specific tension $\left(60 \frac{N}{cm^2} \right)$</p>	Rajagopal et al., 2016 [92] Luis et al., 2022 [266]

$MIF_{gen} \left(\frac{m}{m_{gen}} \right)^{\frac{2}{3}}$	<p>MIF_{gen} = MIF of the generic model</p> <p>m = subject's mass</p> <p>m_{gen} = mass of the generic model</p>	Van der Krogt et al., 2016 [267]
$MIF_{gen} \frac{m}{m_{gen}}$	<p>MIF_{gen} = MIF of the generic model</p> <p>m = subject's mass</p> <p>m_{gen} = mass of the generic model</p>	Modenese et al., 2018 [105]
$\frac{V_M}{OFL} \sigma$	<p>V_M = muscle volume</p> <p>OFL = optimal fibre length</p> <p>σ = specific tension $\left(61 \frac{N}{cm^2} \right)$</p>	van Veen et al., 2019 [268]

3.4. Simulation environment

In this thesis, OpenSim was the software used for all the simulations. OpenSim is an open-source software package that enables users to define, build, and analyse MSK dynamics models, as well as conduct dynamic simulations of movement. It represents a powerful and free tool for the worldwide biomechanical community and serves as a platform for sharing models, results and simulation codes. Based on the application and the available data, two simulation pipelines are implemented in OpenSim: the inverse problem and the forward problem.

In the inverse methods, the data obtained from observed moments represent the starting point to compute quantities involved in generating that movement and that are a function of the model's states (e.g., to predict joint angles and coordinates, joint moments and torques, muscle forces, muscle activity, musculotendon dynamics). Typically, the states of the model include its coordinates, coordinate velocities, muscle activations, and muscle fibre lengths.

On the contrary, the forward problem predicts the observed moments by implementing muscle-driven forward simulations of movement from the neural commands, which include muscle excitations, joint torques, and/or other applied forces.

In the next paragraph, a more detailed description of the Static Optimization tool and how it works is provided since it represents the simulation environment used in this thesis.

3.4.1. Static Optimisation

Static optimisation (SO) is an extension of the inverse dynamics analysis that further solves the net joint moments into individual muscle forces at each instant in time on the basis of prescribed performance criteria, such as minimising the sum of the squared muscle forces. In general, SO solves

the muscle redundancy problem predicting muscle activations and forces by satisfying three main conditions: the joint equilibrium at each instant of time, the tetanic force of each muscle as an upper bound of force a muscle can generate, and the prescribed optimisation criterion. The motion of the model is entirely determined by its generalised positions, velocities, and accelerations. Based on the well-known kinematic, SO resolves the equations of motion, determining the unknown generalised forces. This process is conducted under specific muscle activation-to-force conditions, chosen between the two following conditions:

ideal force generators:

$$\sum_{m=1}^n (a_m F_m^0) r_{m,j} = \tau_j \quad \text{Eqn. 3.5}$$

or constrained by force-length-velocity properties:

$$\sum_{m=1}^n [a_m f(F_m^0, l_m, v_m) r_{m,j}] = \tau_j \quad \text{Eqn. 3.6}$$

while minimising the following objective function implemented in OpenSim:

$$J = \sum_{m=1}^n (a_m)^p \quad \text{Eqn. 3.7}$$

where n is the number of muscles of the model, a_m is the activation level of the muscle m , F_m^0 is the maximum isometric force of the muscle, l_m is its length, and v_m its shortening velocity, $f(F_m^0, l_m, v_m)$ is the force-length-velocity surface; $r_{m,j}$ and τ_j are, respectively, its moment arm and the generalised force acting about the j^{th} joint axis, and p is a user-defined constant (usually set at 2 to implement the optimal control, thus minimising muscle activations). Other objective functions have been implemented, such as the muscle forces [269] or the joint contact forces [268], to better represent the population and the task that users wanted to study.

3.5. Model credibility assessment

Over the last two decades, there has been a proliferation of MSK modelling and simulation in the biomechanics research community. However, the lack of standardised verification and validation practices continues to pose a significant obstacle to large-scale adoption and clinical use. This has spurred intense research activity, which resulted in a few attempts [270–272]. In general, for the credibility assessment of the MSK dynamics models, VVUQ procedures must be followed, and the ASME V&V-40:2018 standard is the reference to pursue credibility. To sum up, model verification focuses on the implementation and numerical accuracy of the model, while model validation assesses

how well the model represents the real world and experimental data. The primary objective of the validation process is to evaluate a computer simulation's ability to predict specific variables of interest by comparing computational results with experimental data. The validation process involves comparing results, extrapolating predictions to intended conditions, and determining the accuracy for the intended use. In musculoskeletal modelling, validation includes aspects like validation experiments and model calibration. However, the level of accuracy required for validation depends on the application rather than the model itself, and it is the users' responsibility to assess whether a model is accurate enough for their intended purpose.

To date, there is no literature on the application of the ASME V&V-40 standard to MSK models, unlike its use in other fields such as computational fluid dynamics modelling [136]. This gap can be attributed primarily to two factors: the lack of standardised procedures for generating these models and the recent introduction of the standard [273]. However, to minimise the influence of the modeler, the ForceLoss project has adopted existing procedures for generating MSK models. These procedures have previously undergone extensive development, rigorous testing, validation, and, in some instances, successful application in clinical contexts. Bones and soft tissues have been manually and semi-automatically segmented by the same expert operator using Mimics v25 (Mimics Innovation Suite, Materialise, Leuven, Belgium), ensuring consistent and replicable segmentation procedures [274]. Subsequently, skeletal models have been generated, and muscle paths were integrated using the nmsBuilder software [101]. This software adheres to ISB recommendations [251] for joint definitions and uses anatomical landmarks for muscle path registration. It has the important advantage of reducing the overall time to generate the MSK models and has been extensively utilised, establishing it as a consolidated protocol [275–277]. Moreover, the definition of the patellofemoral joint aligned with validated techniques found in existing literature, particularly in applications such as gait analysis [275,278,279]. Finally, muscle properties (e.g., OFL and TSL) were morphometrically optimised, utilising a procedure [280] widely adopted within the field [275,281–284]. This approach ensured that the models were constructed upon robust foundations since it leveraged the existing knowledge and methodologies in the biomechanics research community to advance the project's goals.

Chapter 4 – FORCELOSS: A SIMULATION FRAMEWORK FOR THE DIFFERENTIAL DIAGNOSIS OF DYNAPENIA BY MEANS OF SUBJECT-SPECIFIC MUSCULOSKELETAL MODEL

This chapter explores the ForceLoss project, comprising of one study on healthy young volunteers (Part I – *Healthy Volunteers*) and one study on patients waiting for a total knee arthroplasty (Part II – *Osteoarthritic Patients*). The chapter consists of three sections (i.e., material and methods, results and discussion). Since the data collection on the patients’ cohort is still in progress, this chapter will focus on the methods. Building on the information provided in Chapter 3, the methods section of the present chapter will provide further details on the specific implementation of said methods in the ForceLoss studies. The results section will report on the main findings from the study on healthy volunteers. To enhance comprehension and avoid confusion, the reader is referred to Figure 3.2, which shows the experimental protocols and implementations for ForceLoss Part I and Part II – highlighting the key differences..

4.1. Materials and Methods

Upon completion of the study, the full experimental dataset will include anatomical and experimental data from a total of forty subjects. To date, the data collection on the healthy cohort has been completed, while the data collection on the patients’ cohort is still ongoing (see Tables 4.1 and 4.2 for participants demographics). . Both studies were approved by the local Ethical Committee (CE AVEC: 216/2020/Sper/IOR and CE-AVEC 30/2021/Sper/IOR) and have been recorded on the ClinicalTrials registry (Clinical Trials ID: NCT05091502, and ClinicalTrials ID: NCT05795348, for the Part I and Part II, respectively). The entire study was conducted in accordance with the Declaration of Helsinki and each participant gave written informed consent prior to participating in the study. All the subjects were enrolled based on the inclusion/exclusion criteria described in Table 3.1 and each patient was evaluated for health condition using a standard clinical protocol involving a clinical team of an orthopaedic surgeon and a physiatrist.

The experimental protocol of both studies was structured similarly, consisting of two distinct sessions: one for MRI scanning and the other to performing the MVIC test, within a week, ensuring that the MRI accurately reflects the current muscle condition. In the study involving TKA patients, additional tests of clinical interest were collected, further details of which are provided below. From

the collected data, four models per subject were generated, each with increasing levels of personalisation. These models were then used to replicate the experimental tests (MVIC tests) in silico. During each personalised step, a simulation of the MVIC test was conducted, and the predicted MVIC torque was compared with the experimental measurements

ForceLoss - Part I

The demographics of the twenty healthy volunteers enrolled in the ForceLoss - Part I project have been reported in Table 4.1 (Age: 32 ± 6 years old, Height: 168 ± 8 cm, Mass: 65 ± 15 Kg, BMI: 23 ± 3 Kg/m²).

Table 4.1 Subject data: age, height, mass, body mass index (BMI), and physical activity (1 – sedentary; 2 – some physical activity; 3 – regular physical activity and training; 4 – regular hard physical training for competition sports). *Participant excluded from the study.

Subject	Age (years)	Height (cm)	Mass (kg)	BMI (kg/m ²)	Physical activity level	Dominant leg
Female 1	27	170	56	19	3	R
Female 2	26	158	52	21	2	R
Female 3	39	167	53	19	3	R
Female 4	38	166	66	23	3	R
Female 5	33	155	50	21	3	R
Female 6	28	158	55	22	3	R
Female 7	29	160	50	20	3	R
Female 8	26	165	49	18	3	R
Female 9	25	158	52	21	3	R
Female 10*	36	163	57	21	2	R
Male 1	31	178	85	26	2	R
Male 2	27	170	59	20	3	R
Male 3	26	170	81	28	3	R
Male 4	33	177	85	27	3	R
Male 5	31	178	98	30	3	R
Male 6	21	181	78	23	3	R
Male 7	40	173	73	24	3	R
Male 8	39	180	77	23	3	R
Male 9	40	170	60	20	3	R
Male 10	29	170	75	25	3	R

The data from one female adult (i.e., subject 10) were excluded from the study due to her lack of full cooperation, resulting in data that were deemed to be of poor quality.

ForceLoss - Part II

The demographics of the osteoarthritic patients already enrolled are shown in Table 4.2 (Age: 74 ± 4 years old, Height: 168 ± 9 cm, Mass: 72 ± 17 Kg, BMI: 25 ± 5 Kg/m²). The enrolment was conducted by the physiatrist following a consultation or directly by the orthopaedic team responsible for the TKA surgical procedure, ensuring adherence to the defined inclusion and exclusion criteria provided in Table 3.1.

Table 4.2 Patients demographics: sex, age, height, mass, body mass index (BMI), and the examined leg.

Subject	Sex	Age (years)	Height (cm)	Mass (kg)	BMI (kg/m ²)	Examined leg
TKA01	Female	70	165	73	26	R
TKA02	Female	80	165	67	24	R
TKA03	Female	75	167	80	28	R
TKA04	Male	79	180	81	25	R
TKA05	Female	72	162	59	22	R
TKA06	Male	71	184	110	32	L
TKA07	Male	72	168	70	24	L
TKA08	Female	72	150	47	20	L
TKA09	Female	78	158	70	28	R
TKA10	Female	78	154	63	26	R
TKA11	Female	76	175	68	22	R
TKA12	Female	67	170	54	18	R
TKA13	Male	70	176	100	32	L
TKA14	Female	70	178	70	22	L

4.1.1. Magnetic Resonance Imaging

MRI scans were collected using a Discovery MR750w 3.0T scanner (GE Healthcare, Chicago, IL-USA). Full lower limb MRIs, covering the region from L3 vertebra to the toes, were acquired with subjects laying on a bed in supine position. Depending on the patient's height, the scan was performed in four or five overlapping sections to ensure complete coverage of the region of interest, using two specific coils to increase the quality of the images (Figure 4.1). The overlap between consecutive

sections was set to 20 slices (minimum), with further adjustments made for the final segment when necessary due to equipment limitations. The overall acquisition took around 20-30 minutes, yielding four sequences (i.e., water- and fat-suppression, in-phase, and out-of-phase) since the proprietary Dixon sequence was chosen, with a slice thickness of 3.94 mm, a slice increment of 2.0 mm, and a pixel size of 0.4688 x 0.4688 mm.

Figure 4.1 3T MRI scanner at the Rizzoli Orthopaedic Institute, with the patient on the resonance table covered by two proprietary coils to increase the quality of the signal and, thus, the final images.



Image elaboration

The images (i.e., DICOM files) were then imported into Mimics v25 (Mimics Innovation Suite, Materialise, Leuven, Belgium) and merged into a single file. For a comprehensive quantitative analysis, the volumes of bones, muscles, and soft tissues were manually or semi-automatically segmented by a single expert operator as outlined in Table 4.3.

Specifically, the semi-automatic segmentation was employed only for the muscles by means of the Muscle Segmentation Tool, part of the Mimics Innovation Suite license (Materialise, Leuven, Belgium), an atlas-based tool that requires a 2D mask of all muscle tissues as input. The MST functions by comparing the voxels in the mask against those in the available atlases, utilising a vote-based system to determine the appropriate muscle mask assignment for each pixel. The process is executed by dividing into upper and lower sections of the left and right leg (see Figure 4.2).

Table 4.3 Bones and muscles geometries segmented. Bone complexes manually segmented and the list of the muscles was semi-automatically segmented using the Muscle Segmentation Tool, with either low or high manual editing a posteriori.

Bones		Muscle
Group name	Bones contained	Semi-Automatic segmentation (Muscle Segmentation Tool)
pelvis	ilium, pubis, ischium, sacrum	thigh and shank muscles* (n = 37)
femur_r	femur	*rectus femoris, vastus intermedius, vastus lateralis, vastus intermedius were also manually edited starting from the results of the semi-automatic segmentation
patella_r	patella	
tibia_r	tibia, fibula	
talus_r	talus	
calcn_r	calcaneus, cuboid, navicular, lateral/ intermediate/ medial cuneiform, metatarsals	
toes_r	phalanges	

The significant advantage of the muscle segmentation tool is to drastically reduce the time required to segment the individual muscle volumes, compared to a manual segmentation. Despite the muscle segmentation tool yielding good results, manual post-processing remained necessary to refine the 3D volume reconstructions. Particular attention was given to the four knee extensor muscles: rectus femoris (RF), vastus intermedius (VI), vastus lateralis (VL), and vastus medialis (VM) – of primary interest for the ForceLoss study.

Figure 4.2 Coronal view of the results of the muscle segmentation tool applied to the left upper leg (i.e., left thigh).



4.1.2. Maximum voluntary isometric contraction test

The MVIC protocol required subjects to perform a minimum of three contractions in both extension and flexion at a knee flexed at different angles. Each maximal contraction was to be held for

approximately 6 seconds, with a recovery period of at least 90 s between contractions to allow for sufficient rest. If the force measured during the third contraction exceeded the second by at least 10%, additional contractions, up to a fourth or more, were conducted to ensure maximum force output was achieved. The subjects were asked to sit upright on the chair with their arms across their chest throughout all the MVIC testing procedures; they were strapped to the chair to minimise movements of the trunk and the tested leg, allowing the isometry of the test. Throughout each contraction, subjects received verbal encouragement from the operators and real-time visual feedback on the torque being generated to promote maximal effort.

ForceLoss - Part I

For healthy volunteers, the torque data were gathered using a load cell with the sensitivity of 2 mV/V and full scale equal to 100 kg, integrated into the COR1 dynamometer chair (OT Bioelettronica, Italy, Figure 4.3). Each subject performed the knee flexions and extensions with his/her dominant leg at 4 different angles of knee flexion: 45°, 60°, 75°, and 90°. The order of the angles of MVIC test was chosen randomly at the beginning of the experimental protocol.



Figure 4.3 Dynamometer chair OT Bioelettronica used for the MVIC test of the healthy adults.

The MVIC test was preceded by a 10-minute warm-up exercise on the cycle ergometer at zero resistance, followed by a sit-to-stand task, a 6-meter walk, and 8 submaximal contractions, performed both in flexion and extension.

ForceLoss - Part II

All MVIC tests were performed with the participant's dominant leg using the Biodex System 4 PRO

dynamometer (Biodex Medical Systems, Inc., New York, USA, see Figure 4.4). Each subject was asked to perform at least three contractions (leg extension) at two different angles of knee flexion and always in the following order: 75°, and 90°. Where possible, depending on the level of pain and fatigue of the patients, three maximal contractions in flexion with the knee flexed both at 75° and 90° and/or additional maximal extensions in a third configuration (knee flexed at 60°) were performed (see Table 4.4 for the details).



Figure 4.4 Biodex System 4 PRO dynamometer used in the laboratory to acquire torque data with patients.

The subjects were asked to hold the maximal contraction for 6 s maximum, with a resting time between consecutive contractions of 90 s.

Table 4.4 Details of the MVIC tests performed by each subject.

Subject	MVIC					
	Extension			Flexion		
	60°	75°	90°	60°	75°	90°
TKA01	V	V	V	V	V	V
TKA02	V	V	V	-	-	-
TKA03	V	V	V	V	V	V
TKA04	-	V	V	-	-	-
TKA05	V	V	V	-	V	-
TKA06	-	V	V	-	-	-
TKA07	-	V	V	-	-	-
TKA08	V	V	V	V	V	V
TKA09	V	V	V	-	V	V
TKA10	V	V	V	-	V	V
TKA11	-	V	V	-	V	V
TKA12	-	V	V	-	V	V
TKA13	-	V	V	-	-	-
TKA14	-	V	V	-	V	V

Prior to performing the MVIC tests, each patient answered two clinical questionnaires (i.e., the WOMAC³ and KSS⁴). Moreover, BIA and hand-grip tests were acquired as clinical measurements for the quantification of body fat and muscle mass, and the strength of the upper limb muscles, respectively. The warm-up included the same exercises as for healthy adults, with the exception of the cyclo-ergometer.

Torque elaboration

The dynamometry data recorded during the experiment were first converted from V to Nm applying a scaling factor provided by the manufacturers. Then, the torque signals were filtered using a zero-lag 4th order Butterworth low pass filter with a 5 Hz cut-off frequency [285] to remove sudden small bursts due to the noise. In order to find the MVIC torques, the 1st derivative of the torque data was computed and the 1000-ms plateau region [286] with the highest mean was defined as the MVIC torque. The plateau was defined as a sequence of 2000 consecutive samples (1000 ms) where the derivative was equal to 0 (± 0.5).

4.1.3. Electromyography

During the initial setup and the MVIC tests, EMG data were collected from eight primary lower limb muscles involved in the knee extension and flexion. Upon the patient's arrival, an expert operator marked the electrode placement sites, prepared the skin, and placed the electrodes as per the SENIAM guidelines [234]. The muscles recorded included the biceps femoris, semitendinosus, rectus femoris, vastus medialis, and vastus lateralis on the dominant leg, as well as the vastus lateralis and biceps femoris on the non-dominant leg. Bipolar surface electrodes recorded the muscle activity, except during ergometer exercises for healthy volunteers.

EMG elaboration

The EMG signals underwent initial processing with a zero-lag 4th order Butterworth bandpass filter, utilizing cut-off frequencies between 20-500 Hz [233]. RMS envelopes of the EMG signals were then computed using a 500-ms time window [287]. To finalize the process, the EMG signals were normalized against the maximum values recorded during various activities, such as sit-to-stand and MVIC tasks, for each muscle.

4.1.4. Superimposed neuromuscular electrical stimulation (patients cohort only)

Following the MVIC test, the subjects were asked to perform an additional set of maximal contractions (only in extension, with the dynamometer arm fixed at 75°). When the plateau of torque

³ <https://www.fisioscience.it/wp-content/uploads/2021/02/womac-italiano-pdf.pdf>

⁴ https://www.kneesociety.org/assets/docs/Knee-Society-Score-PREOP_ITA.pdf

was reached (approximately 2-3 seconds after onset) a neuromuscular electrical stimulation was delivered to induce the tetanic activation of the vastus medialis. Two reusable wetted rubber electrodes were placed medially on the anterior aspect of the upper thigh and on the anterior aspect of the lower thigh (the anode and the cathode, respectively), to stimulate a representative (i.e., greater than 20%) sample of the muscle belly. A doublet of single square-wave stimuli was delivered through a constant current high voltage stimulator (Digitimer DS7AH, Hertfordshire, UK). The stimulus consisted of 2 square pulses, each with a 100- μ s duration, 300 V maximal voltage, and intensity adjusted from 200 to 500 mA, with a 10 ms interpulse interval. The intensity of the stimuli was selected based on the stimulation required to reach at least 25% of the peak torque previously recorded. A resting period of 3 minutes was granted between contractions to avoid muscle fatigue, and to allow for a full recovery after the stimulation. The Power1401 data acquisition system (CED, Cambridge, UK) was used for an adequate synchronization between dynamometry data and SNMES, that were visualized and recorded in Spike II v10 (Cambridge Electronic Designed Limited-CED, Cambridge, UK).

4.1.5. BIA (patients cohort only)

All the patients were asked to lie down on a physiotherapy bed. Electrodes were respectively placed on a hand and foot, using a foot-to-hand bioimpedance technology at 50 kHz frequency. The whole-body BIA was assessed using the BIA 101 BIVA Pro device (Akern Systems, Firenze, Italy), a small machine to which the electrodes were connected. The device gave as outputs the resistance (R_z) and reactance (X_c), and the phase angle (θ). Using equations provided by the vendor, the raw measures were employed to estimate the fat-free mass (FFM), the fat mass, and the percentage of FFM.

4.1.6. Hand-grip test (patients cohort only)

Following the BIA analysis, the patients were asked to perform an hand-grip test. The test consisted of the repetition of at least three maximal contractions interspersed by 90 s of rest time while holding a hydraulic dynamometer (Jamar) in the dominant hand, with the elbow flexed at 90°. In case the maximal contraction force value (in kg) was scored at the third contraction, a fourth attempt was requested. Throughout all the tests, verbal encouragement was provided to elicit maximal contractions.

4.1.7. Modelling pipeline

Subject-specific models have been generated starting from the segmentations of bones, muscles and soft tissues, previously segmented in Mimics v25 (Mimics Innovation Suite, Materialise, Leuven, Belgium), following the pipeline proposed by Modenese and colleagues [105]. In nmsBuilder software [101], a MSK model for each subject (i.e., both volunteer and patient) was created, each comprising of seven segments (pelvis, thigh, patella, shank, talus, calcaneus, and toes), 13 degrees of

freedom and 40 muscles. Each model's kinematic chain was crafted by fitting analytical shapes (i.e., cylinders or spheres) to the bones' articular surfaces, establishing joint reference systems necessary for defining the joint centres and orientation of axes, using the MeshLab software [288]. Moreover, patellar movements were modelled to capture the 3D motion with a single degree of freedom dependent to the knee joint, ensuring accurate representation of joint mechanics. The three-dimensional motion of the patella during flexion-extension was enabled by parameterising splines, which were adapted from existing literature [89,92]. These splines were scaled to each patient based on palpated anatomical landmarks and adjusted to ensure that the patella correctly articulated with the femur [278]. In nmsBuilder, the inertial properties of each segment were also defined (assuming the densities equal to 1.42 g/cm³ and 1.02 g/cm³, respectively for bones and soft tissues [289,290]), and muscle-tendon units were personalised based on individual anatomical data. Muscles were modelled as Hill-type actuators [291], described by Millard et al. [258], and their behaviour was adjusted to prevent unrealistic penetrations into bone, using wrapping geometries when necessary. Each muscle's line of action was checked against the segmented muscle volumes, with minimal manual adjustments made where required. Muscle parameters such as pennation angle, optimal fibre length (OFL), and tendon slack length (TSL) were initially derived from the generic Full Body model [92] and then anthropometrically scaled to fit each subject's specific dimensions [280]. The maximal isometric force (MIF) for each muscle was recalculated differently based on the level of personalisation of the MSK dynamics model.

Personalisation of Muscle Properties

Based on the personalisation of two muscle properties (i.e., the MIF and the activation level), four different subject-specific models of each subject were generated.

Initially, the base model, referred to here as M_{genPCSA} , utilised MIF values of a generic healthy age-matched population, scaled according to a linear regression model [265], based on the height and mass of the subjects being studied. Subsequent enhancements were made by incorporating experimentally collected and processed data, leading to the development of three additional models. The M_{ssPCSA} model incorporates subject-specific PCSA obtained from MRI scans for defining MIF values according to the following equations (Eqn. 4.1) and (Eqn. 4.2):

$$PCSA_i = V_m^i \cos \theta / l_o \quad \text{Eqn. 4.1}$$

$$MIF_i = \sigma PCSA_i \quad \text{Eqn. 4.2}$$

where V_m is the i -th muscle volume, θ is the pennation angle, l_o the OFL and σ the specific tension, set 60 N/cm^2 and in line with literature data [92,265]. The $M_{ssPCSAAssEMG}$ model was based on the M_{ssPCSA} model, with the inclusion of adjusted maximal control settings for the quadriceps muscles (RF, VI, VL, VM) based on EMG data, deviating from the default maximal activation value of 1. The last model (i.e., $M_{genPCSAAssEMG}$) built upon the $M_{genPCSA}$ framework but integrated EMG-based maximal activations similar to the adjustments seen in the $M_{ssPCSAAssEMG}$ model. For the latter two models, the peak values of the normalised RMS envelopes for the knee extensors were utilised to regulate the quadriceps activation levels to mirror experimental data accurately. Where experimental data were lacking, specifically for the VI muscle, activation levels were inferred from the mean activations of the VL and VM muscles [292,293], reflecting the synchronised activity pattern commonly seen among the three vastii muscles [294–297]. The varying degrees of personalisation across these MSK models are detailed in Table 4.5.

Table 4.5 Summary of the different personalisation steps of the models (ss = subject-specific).

	ss bones	ss muscle volume/force	ss muscle activation level
$M_{genPCSA}$	V	X	X
M_{ssPCSA}	V	V	X
$M_{ssPCSAAssEMG}$	V	V	V
$M_{genPCSAAssEMG}$	V	X	V

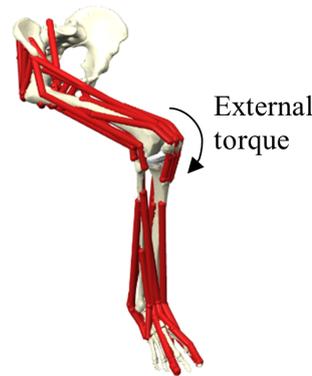
4.1.8. Biomechanical simulations

Each model was placed in a seated position, with the hip and the knee flexed at 80° and 90° , respectively, to mimic the experimental setup. The boundary conditions (i.e., human-seat interactions) were simplified by locking the pelvis translations, not allowing the pelvis body to move in space throughout the simulation. This is a common approach [257,298] to simulate a maximal isometric leg extension task when the forces transferred from the seat to the human body are unknown due to the lack of experimental data. Nonetheless, it ensures the model is dynamically consistent.

In OpenSim [103], once the boundary conditions were defined, a Static Optimization (SO) simulation was run hypothesising optimal muscle control (i.e., minimising the sum of squared muscle activations [299]) while (i) imposing a static kinematics (i.e., the degrees of freedoms in the model were kept constant throughout the whole simulation) and (ii) applying a constant flexion torque at the knee joint (Figure 4.5). The external torque was iteratively increased in 1 Nm steps until the model was unable

to sustain it (i.e., due to quadriceps weakness). The last torque value prior to failure was considered to be the model's MVIC torque.

Figure 4.5 Imposed boundary condition in the MVIC simulation.



4.1.9. Simulation environment

The core of the ForceLoss solution is the simulation framework, which was specifically devised to support the differential diagnosis for dynapenia. In this, personalised MSK dynamics models serve as falsification tools to test different clinical hypotheses, allowing to discard the less plausible ones (i.e., those that cannot explain the experimentally observed loss of muscle force) and to identify the potential primary cause of dynapenia. More specifically, the framework is developed in multiple stages (referred to Figure 3.1), to exploit the collected data fully and gradually. At each step, the MSK dynamics models are progressively personalised and subsequently employed to replicate - in silico - the MVIC test. The maximal torques predicted by the model are then compared to the experimental values measured by the dynamometer to reach an informed decision. However, prior to deploying the comprehensive ForceLoss framework, an essential preliminary study was conducted utilising data from healthy subjects. This initial investigation is essential for evaluating the accuracy with which models could be constructed and the MVIC test could be replicated in silico, given that complete data on elderly subjects were not initially available and still remain incomplete.

For each healthy volunteer, four subject-specific MSK dynamics models were generated with increasing level of personalisation and, at each step, used to simulate a computational MVIC test. Initially, a base model was created using average PCSA values derived from healthy adult data ($M_{genPCSA}$), with muscles allowed to reach tetanic contraction (maximal activation level set at 1) to predict the MVIC torque. If the predicted MVIC torque significantly exceeded the measured value, the simulation was repeated using the M_{ssPCSA} model, which incorporates subject-specific PCSA values. If the model's prediction continued to exceed the experimental value, even after incorporating

patient-specific PCSA values, EMG data from the MVIC test were used to adjust for non-pathological submaximal activation levels ($M_{ssPCSA_{assEMG}}$). Lastly, a fourth model (i.e., $M_{genPCSA_{assEMG}}$) was developed to assess if personalised muscle activation levels alone, without using subject-specific muscle volume segmentations, could accurately simulate the experimental MVIC torque.

The ForceLoss simulation framework

The framework starts once the diagnosis of dynapenia has been confirmed, i.e., if the measured maximal torque significantly falls below the normative ranges reported for an age-matched healthy population.

An initial image-based model is developed using the available MRI data. This base model ($M_{genPCSA}$) features MIF values derived from data from an age-matched healthy population; thus, it does not account for the actual muscle size/volume. As is, the model – de facto representative of a healthy individual of the size of the subject under study – should, in principle, overestimate the experimental values.

At this point, the MIF values in the model are personalised based on the muscle volumes segmented on the medical images. The eventual presence of sarcopenia is thus accounted for. The updated model (M_{ssPCSA}) is employed to perform the *in silico* MVIC test. If the predicted torques approximate the experimental findings well, sarcopenia is reasonably assumed to be the entire cause of dynapenia. Otherwise, alternative clinical scenarios are explored.

First, the experimental torque is compared to the maximal torque observed while delivering neuromuscular electrical stimulation (i.e., during the involuntary contraction). In case of disagreement, if the predicted torque is higher than the voluntary contraction torque, but comparable to that obtained by forced involuntary contraction, then the dynapenia is attributed to muscle activation inhibition.

Finally, if the model so informed predicts an MVIC torque that exceeds the experimentally measured torque and from the EMG data reveals atypical muscle activations, an EMG-assisted simulation is conducted, thus verifying the potentially suboptimal motor control as a cause underlying the dynapenia.

4.2. Results

ForceLoss - Part I

In the following, preliminary results using data from the healthy cohort are presented. In particular, the results obtained from the data of the nine healthy young females showed that introducing a second level of personalisation using MRI or EMG data decreased the mean relative error from about 25%

with the base model (M_{genPCSA}) to roughly 15%, with no significant differences in performance between the M_{ssPCSA} and $M_{\text{genPCSAAssEMG}}$ models. Further reductions in relative error to 5% were achieved when both anthropometric measurements and subject-specific muscle activations (the $M_{\text{ssPCSAAssEMG}}$ model) were employed.

On an individual basis (Figure 4.6), the M_{genPCSA} model accurately reflected experimental data for only one (of nine) subjects (Figure 4.6.A, subject 7). Refinements using segmented muscle volumes to adjust the MIF values significantly lowered the simulated extension torque, aligning it more closely with the measured values. Of note, for subjects 6 and 8, the M_{ssPCSA} model's predictions were less accurate than those of the M_{genPCSA} model (respectively, 169 vs 162 Nm, and 187 vs 174 Nm compared to the real value of 146 Nm and 127 Nm). Figure 4.6.C). Incorporating EMG data to specifically limit the maximum activation levels of the quadriceps muscle ($M_{\text{ssPCSAAssEMG}}$) significantly improved model predictions, lowering the prediction error to 6% or less – on average - for all nine subjects, demonstrating a substantial enhancement over previous models.

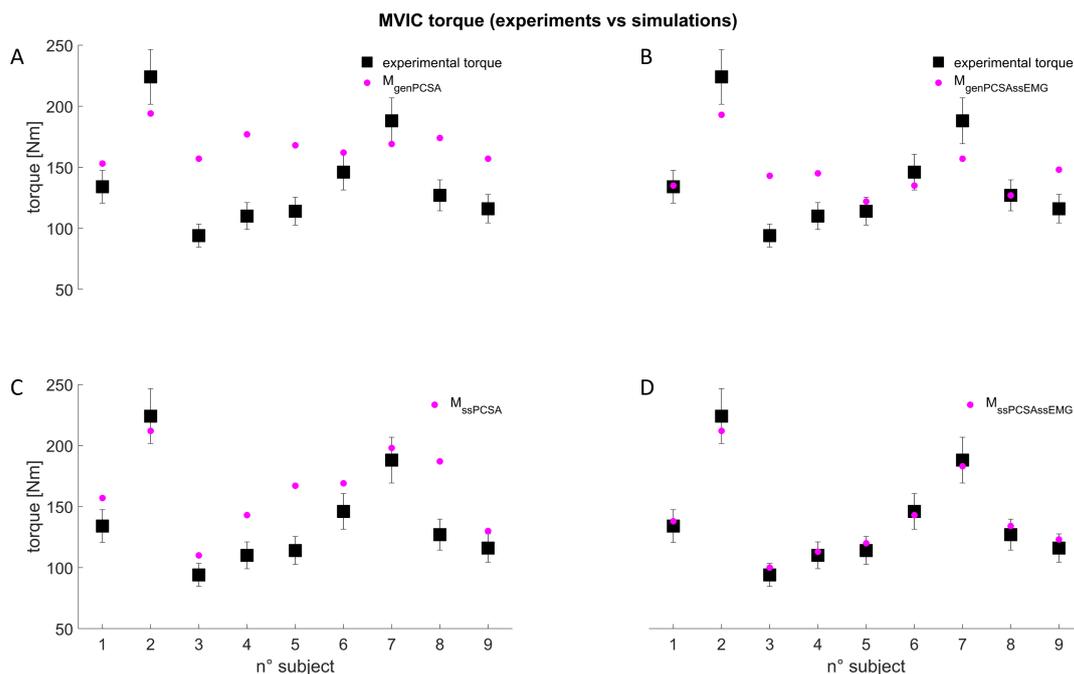


Figure 4.6 The one-to-one comparison between the maximal torque predicted by the simulation (purple dot) and the real data (black square) is provided for the nine healthy female subjects using the different models: (A) M_{genPCSA} , (B) $M_{\text{genPCSAAssEMG}}$, (C) M_{ssPCSA} , and (D) $M_{\text{ssPCSAAssEMG}}$.

While the use of $M_{\text{genPCSAAssEMG}}$ models allowed to minimise the mean error across participants, for four subjects the difference between experimental and simulated values remained lower than the 10% error threshold (135 vs 134 Nm, 122 vs 114 Nm, 135 vs 146 Nm, 127 vs 127 Nm, respectively for subject 1, 5, 6 and 8).

ForceLoss - Part II

As mentioned at the start of this chapter, the acquisition of patient data is still ongoing and the data processing has just begun. Therefore, the only findings shown relate to the data acquisition. All the data included in the data collection protocol that have been recorded for the first fourteen enrolled patients are summarised in Table 4.5.

Table 4.5 Clinical measurements and data acquired in the experimental protocol for each patient, including medical images, isometric dynamometry test, involuntary contractions, bio-impedance analysis and had-grip dynamometry.

Subject	MRI	MVIC	SNMES	BIA	Hand-grip (kg)
TKA01	V	V	V	-	V
TKA02	V	V	V	-	V
TKA03	V	V	V	-	V
TKA04	V	V	V	V	V
TKA05	-	V	-	V	V
TKA06	V	V	V	V	V
TKA07	V	V	V	V	V
TKA08	V	V	-	V	V
TKA09	V	V	V	V	V
TKA10	V	V	V	V	V
TKA11	V	V	V	V	V
TKA12	V	V	V	V	V
TKA13	V	V	-	V	V
TKA14	V	V	V	V	V

4.3. Discussion

The primary objective of the ForceLoss project is to demonstrate the feasibility of a framework for the differential diagnosis of dynapenia in orthopaedic patients who are candidates for TKA. However, prior to realising this goal, it is essential that both the MSK model and the simulation framework undergo rigorous testing and validation.

In the initial phase, data from healthy volunteers were employed to set up and test the simulation environment where to conduct an MVIC test *in silico*, integrating experimental data (e.g., EMG recordings, dynamometry data, and medical images) and computational tools (i.e., personalised MSK models). This phase also involved a preliminary validation wherein the simulated outcomes were compared against experimental measurements. Additionally, this study sought to determine the minimal level of personalisation necessary to achieve accurate enough predictions, i.e. within 10% of the experimentally measured values. For each of the nine healthy young female volunteers, four MSK models with varying degrees of personalisation were generated using MRI data. The muscle properties within these models were tailored based on several criteria: generic muscle volumes representative of an age-matched population (M_{genPCSA}), PCSAs derived from medical imaging data

(M_{ssPCSA}), muscle maximum activation levels ($M_{genPCSAAssEMG}$) based on the available surface EMG data, or a combination of PCSAs and muscle maximum activation levels ($M_{ssPCSAAssEMG}$).

Our preliminary findings revealed that digital twins could predict maximal knee extension torques close to the in vivo experimental MVIC torque values observed in a healthy female cohort. However, the sole use of anthropometric measurements derived from medical images was not sufficient to achieve acceptable prediction errors (i.e., $<10\%$). To meet the set threshold of accuracy personalised muscle activation levels (i.e., $M_{ssPCSAAssEMG}$) were to be enforced. It is important to note that the experimental MVIC torques (measured with the dynamometer arm set at 90°) were in agreement with previous literature for an age-matched population [300–302]. Nonetheless, when using population data to scale the PCSAs, the models tended to overestimate the experimental values. By incorporating personalised information from MRI data to adjust the MIF values (M_{ssPCSA}), the prediction error was generally reduced (compared to the results from the $M_{genPCSA}$ models), and for three subjects it fell within the predefined 10% threshold. However, only the combined personalisation of both MIF values and maximal activation levels (extracted from experimental EMG data) enabled to achieve low errors for all participants.

The entire study is not devoid of limitations. Firstly, the lack of kinematic data to accurately depict knee joint motion during the MVIC test necessitated maintaining a static pose throughout the simulation. Secondly, the inability to directly measure certain muscle parameters such as OFL and TSL from the available data (i.e., proprietary Dixon medical images) [303] compelled the use of alternative computational methods, as ensuring that the force-length-velocity relationships of the generic Full Body Model and the ratios between OFL and TSL were preserved [280]. Furthermore, one notable constraint is the reliance on images from high-cost 3T MRI equipment, which incurs significant expenses. However, recent advancements in technology suggest that the estimation of muscular geometries could be achieved using more accessible 3D ultrasound technologies [304]. This shift could potentially reduce both the computational costs and time required for the proposed pipeline. Additionally, recent advancements in motor unit (MU)-driven musculoskeletal modelling now enable the simulation of peak firing rates across complete populations of MUs within a muscle, providing more accurate estimates of muscle-tendon force [305,306]. These developments suggest that incorporating newer emerging technologies could significantly enhance the effectiveness and efficiency of musculoskeletal modelling and, therefore, the ForceLoss project, making it a more viable option for widespread clinical application.

The next step will be to develop subject-specific MSK dynamics models for each of the elderly participants enrolled in the ForceLoss – Part II study (similarly to what it is hereby presented for a

cohort healthy young adult females), and to use them to perform a MVIC test in silico. If the experimental data will highlight a reduced muscle force, compared to normative data from an age-matched healthy population, the first simulation (employing M_{genPCSA}) will be used to confirm the diagnosis of dynapenia. More specifically, if the predicted values will be significantly larger than the corresponding experimental data, the subject will be considered by all means dynapenic. The model will be then personalised to account for subject-specific muscle volumes; if the prediction will be in good agreement with the experimental value, the dynapenia will be assumed to be primarily caused by sarcopenia. If the values will still differ, the actual tetanic force of the knee extensors will be reduced to account for submaximal activation. In case this level of personalisation did not produce good agreement between the model and experiment, further analysis of the EMG signals will be conducted to investigate possible muscle coordination problems. Moreover, to take into account possible problems in the motor control, the framework will be complemented with two additional steps, where experimental superimposed neuromuscular electrical stimulation data and EMG-assisted simulations will come into play.

For elderly subjects, the limitations previously noted can be further compounded by additional challenges inherent in assessing this demographic. Despite the exclusion criteria, due to the nature of aging, it is possible that in some cases, subjects may not be fully suitable for all tests. Consequently, these patients may not undergo all planned measurements or analyses. For example, one patient (i.e., TKA05) could not undergo subject-specific modelling due to a metallic pin that interfered with MRI scans. Additionally, cardiovascular issues prevented the use of electrical stimulation for some patients, which significantly limited the evaluation options available for confirming muscle function and diagnosing conditions accurately. Moreover, the limited availability of materials necessary for conducting bio-impedance analysis further complicated data collection. The onset of fatigue also markedly affected compliance with the MVIC protocol among patients, underscoring the important role of motivation in tasks requiring maximal voluntary contractions (and, thus, without neuromuscular electrical stimulation). These instances highlight the complex considerations involved in conducting research with elderly populations, where adaptations may be necessary to accommodate individual health statuses while striving to achieve research objectives.

Chapter 5 – GENERAL DISCUSSION AND CONCLUSION

The beginning of the last century has witnessed remarkable advances in technologies employed in the study of human movements. These technological innovations, including wearable sensors and computational simulations, have revolutionised the study of human mobility, offering unprecedented insights into biomechanics and enhancing our understanding of various physiological processes. Wearable sensors have been demonstrated to be capable of providing researchers with data acquired continuously over long periods of time and in the real world while people perform their daily activities, thus allowing the quantification of mobility performance. The widespread adoption of this technology has been driven by the great advances in device miniaturisation and the development of a new generation of algorithms that enable accurate and reliable measurements of mobility-related parameters in subjects with slow or pathological gait.

In parallel, the rise of *in silico* medicine has led to a paradigm shift in the analysis of human biomechanics. Computer models, especially musculoskeletal dynamics models, allow researchers to simulate and predict the behaviour of the human body under various conditions or computing properties (e.g., internal joint contact forces) that are difficult or even impossible to measure without resorting to invasive procedures. The application of computational simulations extends beyond basic biomechanics, with potential implications for clinical practice. *In silico* medicine facilitates the development of personalised treatment plans by simulating the effects of interventions, predicting outcomes, and optimising therapeutic strategies. For instance, orthopaedic surgeons can use the outcomes of the simulations to plan surgeries, assess the impact of different procedures, and tailor interventions to individual patients.

However, the interpretation of the results obtained with the above technologies demands a deep understanding of their limitations and potential sources of error. Before these technologies can be used, their credibility must be assessed, specifically when employed to interpret, diagnose, and predict treatments or diseases. Indeed, the credibility assessment of such methodologies remains a challenging and complex endeavour, characterised by its time-consuming, expensive, and occasionally ambiguous nature. Validating the accuracy and reliability of wearable sensors and computational models requires rigorous testing against established gold standards.

In this context, the focus of this PhD project was to investigate the regulatory pathway for the credibility assessment of different models: the software pipeline to elaborate IMU data and extract digital mobility outcomes, and subject-specific (neuro)musculoskeletal models. After briefly

outlining how the knowledge gained from the Mobilise-D project has been and could be applied to the ForceLoss project to simplify the narrative and enhance readability and comprehension, the discussion will be structured into two sub-sections.

The bridge between Mobilise-D and ForceLoss projects

In Mobilise-D, interactions with regulatory agencies, including both the EMA and the FDA, were essential to gain insights into the procedures required for a credibility assessment aimed at obtaining marketing authorization for a medical device, specifically the Mobilise-D analytics software (combined with a wearable sensor) to extract mobility-related parameters. The knowledge derived from these interactions was subsequently leveraged in the ForceLoss project. Based on the FDA's response to the Pre-submission request, it has been confirmed that the ASME V&V-40 standard can indeed be applied to MSK models. Moreover, looking further, the ForceLoss project could be potentially used as a stratification tool during the patient enrolment phase of drug trials (i.e., leveraging MSK models to differentiate dynapenic patients and identify those who exhibit sarcopenia without other concurrent factors). In this regard, the information gained from interactions with the EMA and FDA during the Mobilise-D project could be valuable for qualifying the ForceLoss methodology. Specifically, the knowledge acquired during the qualification advices with the EMA for Mobilise-D's novel methodology for drug development could be applicable to ForceLoss.

5.1. On the Mobilise-D project

The rapid proliferation of wearable sensor technologies and, in particular, the concurrent rapid evolution of data processing algorithms, not solely based on physics laws anymore but with more numerous data-driven components, represents a significant facet of contemporary research. Notwithstanding, the regulatory authorities have struggled to keep pace, raising concerns about ethical implications, data privacy, and the potential for misuse.

In this regard, this PhD thesis tried to delve into the transformative impact of these advancements, emphasising the shift from traditional physics-based algorithms to more adaptive, data-driven methodologies. While these developments hold immense promise for enhancing our understanding of biomechanics and human movements, a critical observation emerges regarding the regulatory framework surrounding these innovations. The metrological approach is the only approach currently accepted to demonstrate the credibility of such a methodology, but it is not exempt from limitations and disadvantages. The Mobilise-D technical validation study, for example, was a huge and expensive endeavour that involved the enrolment of patients in five sites across Europe and Israel. Moreover, in case of any changes in the software, like an update or a modification, a new study must be designed and conducted, with obvious feasibility (and economical) issues that make it virtually impossible for

any manufacturer to afford it. Yet, there is a lack of globally recognised alternatives for assessing the credibility and, thus, obtaining regulatory approval, which calls for a harmonised and agile regulatory response that aligns with the dynamic nature of recent technologies. The current discussion must address the need for regulatory frameworks that can adapt to the dynamic nature of the advancements in the software component of wearable sensors, ensuring a balanced integration that upholds ethical standards in the face of rapidly evolving technological landscapes. The Mobilise-D project has highlighted that the marked distinction between metrological and verification and validation (e.g., the risk-based process developed in the ASME V&V-40:2018) approaches needs to be overcome with the aim of reducing the time and effort for the certification procedure.

Future developments

In the regulatory context, a structured report that functions as a guideline for industries navigating the regulatory pathway for technology qualification approval, similar to the one proposed by the Mobilise-D project, could prove useful. It would serve as a comprehensive resource offering insights into the qualification process, aligning with regulatory standards. The two letters of support published by the EMA in response to the request for qualification advice submitted by the Mobilise-D consortium already encapsulate the collaborative efforts and regulatory milestones achieved with the EMA. Furthermore, ongoing efforts are underway to enhance transparency and share crucial insights into the regulatory landscape highlighted in the feedback from the FDA staff from the CDRH in response to a pre-submission form. This submission sought to explore and determine the optimal regulatory process for certifying the Mobilise-D analytics software, investigating alternative certification pathways to overcome the limitations of the currently approved methodologies. Sharing all the different feedbacks received in the field of the regulatory certification from various regulatory agencies (i.e., both the EMA and the FDA) could represent a rigorous and well-informed document of reference for the regulatory pathways. This comprehensive report could be a valuable resource for the manufacturers or the industries, providing important information to ensure alignment with international regulatory standards and fostering a more collaborative and informed approach to the technology's qualification approval process.

5.2. On the ForceLoss project

Overall, the ForceLoss project sought to enable the differential diagnosis of dynapenia in elderly subjects waiting for a total knee arthroplasty through a comprehensive approach, combining experimental measurements such as medical images, electromyography, dynamometry tests, bioelectrical impedance analysis, and neuromuscular electrical stimulation with computer models and simulations. The project was structured into two parts, driven by both organisational considerations

and differences in the population enrolled. The ForceLoss: Part I study aimed to define, test and validate a new simulation framework devised to support the differential diagnosis of dynapenia using the data collected on twenty healthy adults, representing the reference dataset. The second part of the ForceLoss project is the more clinically oriented part as it involves the enrolment of twenty subjects, scheduled for primary total knee arthroplasty surgery and affected by dynapenia. The main objective of ForceLoss: Part II is to demonstrate the applicability of the framework, with the addition of steps to test different (clinical) factors possibly contributing to dynapenia.

In the last two decades, the use of musculoskeletal modelling and simulations in clinical settings (e.g., to plan surgeries or treatments, or as tools to support clinical decision-making) has increased, opening up new perspectives in healthcare. However, the credibility assessment to evaluate the level of reliability of the model's predictions still represents a bottleneck for large-scale use, primarily due to the limited reproducibility of models. Ideally, a model based on the same data should yield identical outcomes, suggesting that the modelling and simulation strategy does not influence the results. Contrarily, evidence [307,308] shows that due to the absence of standardised procedures universally recognised as valid, the personal choices of each modeller, influenced by their prior knowledge and experience, significantly affect the model generation, leading to highly variable results. Regulatory agencies are aware of these challenges and recognise computational modelling as a regulatory science priority [309]. From a practical standpoint, the recent introduction of the ASME V&V-40 standard supports the advancement of credibility methods and exemplifies efforts to bring rigor and transparency to computational modelling. This standard, which provides a structured approach for model verification and validation, has proven useful in enhancing model credibility, particularly in finite element models [136]. Nonetheless, due to the novelty of this standard, its application to musculoskeletal modelling has not yet been documented in the literature. Musculoskeletal models, crucial for biomechanical analysis and medical diagnostics, would greatly benefit from such standardised validation processes. The implementation of these standards can help overcome current challenges by providing a methodological framework for evaluating the accuracy and appropriateness of model choices, thereby enhancing their trustworthiness and clinical applicability. As the field progresses, it is crucial that these computational tools undergo rigorous testing and validation, following protocols similar to those in other scientific fields, to ensure their efficacy and safety in clinical decision-making [273,307].

However, hereafter it is provided a possible way to apply the ASME V&V-40, comprehensive of a broader validation, to the MSK models as defined in the ForceLoss project.

Hypothetical application of the ASME V&V-40 to the MSK model

The ASME V&V-40 framework includes four main steps (Table 5.1). Initially, it requires the definition of the question of interest, specifying the problem or decision scenario where the model will be applied. Following this, the COU is detailed, outlining the model’s scope and its specific role in addressing the question. The process continues with an assessment of model risk, which involves evaluating the potential for the model’s use to result in decisions that could cause patient harm or other undesirable outcomes. This risk is viewed as a combination of the model’s impact on decision-making and the consequences of any adverse outcomes due to incorrect decisions based on the model. Finally, credibility factors are established for various elements of the verification and validation process. These goals, driven by the risk analysis, ensure that the model is both safe and effective for its intended use [136].

Table 5.1 Hypothetical question of interest, context of use, model influence and decision consequence for the ForceLoss project.

Question of interest	Can the MSK model accurately predict the MVIC torque of the knee extensors?
Context of use	MSK model to support the differential diagnosis of dynapenia
Model Influence	<i>Low</i> → the model influence is low because the differential diagnosis of dynapenia is based on the clinicians’ diagnosis and just supported by the model
Decision consequence	<i>Low/Medium</i> → if the model wrongly influences the clinicians’ decision, the disease progresses and patient’s health status does not get better.
Model risk	<i>Low-Medium</i> (Level 2 out of 5)

Based on the model risk, the credibility factors and the rigor required must be defined (low-medium level as shown in Table 5.1). For a full and comprehensive credibility assessment all the thirteen credibility factors, as specified within the ASME VV40 standard, should be evaluated. For instance, the initial selection for the rigor level of the credibility factor for the validation assessment (i.e., *output comparator*) was defined as *Comparison performed by determining the difference between computational results and experimental results*. Nonetheless, it would be more suitable to adopt a higher accuracy level, specifically *Uncertainty in the output of the computational model or the comparator was used in the output comparison* [136], to improve the precision and dependability of the evaluation.

Limitations

Several limitations were encountered in the ForceLoss project. The lack of kinematic data to accurately describe knee joint motion during the execution of the MVIC test compelled the use of a completely static pose while simulating the test. This choice was driven (justified) by the isometric nature of the experimental test. However, although the participants were kept in position by means of seatbelts and straps throughout the execution of the entire test, (minimal) leg motion could be observed. Moreover, in the definition of important muscle parameters, such as optimal fibre length and tendon slack length, it was necessary to resort to the use of morphometrically scaled generic values derived from the Full Body model. In vivo measurements of the subject-specific musculotendon properties are difficult to accurately estimate. The mentioned limitations are related to the modelling and simulations and, therefore, are shared by both parts of the ForceLoss project. Moreover, significant limitations arose in the experimental protocol with the osteoarthritis patients, as, for various reasons, the dynapenic subjects were unable to carry out the experimental protocol in its entirety. The majority of the patients experienced difficulties in performing flexions and due to fatigue onset, combined with the ageing and the clinical status, reduced the MVIC protocol to two angles of knee flexion (75° and 90°) and mostly to extension only. The common perception of feeling unmotivated is noteworthy, particularly since maximal contraction without neuromuscular electrical stimulation relies completely on voluntary effort, coupled with the emotional stress of having to undergo a surgery. Finally, unprecedented inconveniences occurred due to the Covid-19 pandemic, including delays in the delivery of instrumentations (e.g., the Biodex dynamometer delivered with a delay of almost two years) or the highly limited use of the hospital clinics for research that caused a consistent shift of the enrolment, and affected the performance of all subjects, who were asked to wear a face mask while performing the MVIC task.

The latter issue highlighted the need for open-access datasets, which could be exploited or combined to existing datasets to develop and test new computational modelling approaches, and – by extension – to advance the biomechanical field. Despite a noticeable trend, over the past few years, to share not only the results of a research but also the associated datasets, some more efforts in this sense are still required. The open-source data availability would facilitate the establishment of large normative databases, to ease the generation and validation of models in different clinical settings. The ForceLoss project will contribute in this regard, since the project fits into the wider In Silico World project, which aims to accelerate the adoption of in silico methodologies by lowering barriers related to the development and validation of model, encouraging and promoting the sharing of data.

An additional limitation of these models is the absence of a comprehensive sensitivity analysis specifically focused on patellar positioning. Variations in the location of the patella directly impact the biomechanical outputs of the model, which could lead to inaccuracies in the simulation results. In this work, to limit the modeller's influence, the patellofemoral joint was defined employing existing techniques already implemented in the field [92,275,310]. This approach, aimed at minimising the variability introduced by the personal choices of the modeller, is indicative of a broader methodology that was consistently applied across all other steps of the modelling process, ensuring uniformity and rigour in the development of the musculoskeletal model. Both bones and soft tissues were segmented manually and semi-automatically using Mimics v25 (Mimics Innovation Suite, Materialise, Leuven, Belgium) by the same expert operator, ensuring consistent and replicable segmentation procedures. This was followed by the generation of skeletal models and the integration of muscle paths using NMSBuilder software, which aligns with ISB recommendations for joint definitions and uses anatomical landmarks for muscle path registration. Muscle properties such as the length of both the fibres and the tendon were morphometrically optimised using a widely accepted procedure, grounding the model development in established biomechanical research methodologies. These foundational practices have enhanced the project's validity, with initial validation efforts in the ForceLoss project comparing the experimentally measured maximal torque against the simulated outputs.

Future developments

The ForceLoss project is still ongoing, and thus, the application of the proposed framework to support the differential diagnosis of dynapenia must be validated on a larger cohort (i.e., including also the ten healthy male adults) and tested on the twenty osteoarthritic patients.

For the ForceLoss: Part I – Healthy Volunteers project, the results obtained with the data of the ten healthy female adults that led to a good agreement between simulation outcomes and the experimentally measured maximal torque must be confirmed using the models of the ten healthy male adults.

The ForceLoss: Part II – Osteoarthritic Patients study is still at its origin, with the enrolment and the data collection still ongoing. The future works will be to test and verify the protocol defined to support the differential diagnosis of dynapenia on all the twenty dynapenic and osteoarthritic patients.

To sum up, the advances in miniature technologies and in silico medicine have ushered in a new era of understanding human movements. Wearable sensors have been proven to enable unprecedented insights into mobility, allowing the quantification of mobility performance. In addition, computer models and simulations are a promising and increasing tool for research and applications in clinical practice. However, the credibility assessment of these methodologies remains a critical yet

challenging aspect. Overcoming the barriers to validation, standardising protocols that have not been updated to keep pace with these advances, and addressing the limitations of these technologies are essential steps towards ensuring their reliability and widespread adoption in the study and application of human movement analysis.

Appendix A – Description of the Mobilise-D analytics software

The Mobilise-D analytics software was implemented to be the optimal analytical pipeline for the quantification of DMOs (e.g., step detection, real-world walking speed) from raw IMU data (Figure 2.2).

Pre-processing

The input to the entire pipeline is the raw IMU data, which are pre-processed to increase the quality of the signal and, thus, to reduce noise, as follows:

1. IMU orientation is estimated with a 6D-IMU (3D accelerometer, 3D gyroscope) complementary filter, allowing alignment of the IMU sensor with the vertical gravity vector [307]. Additionally, the principal component analysis technique is used to align the IMU signals with the direction of the body movement [308].
2. Detrending, low/high/band-pass filters with various cut-off frequencies, continuous wavelet transform and 1D morphological filters [309] are applied in the pre-processing step.

Gait sequence detection

Once processed, the IMU data enters the gait sequence detection (GSD) block, where it is analysed to identify the gait sequences (defined as a repetitive pattern delimited by two consecutive contacts of the same foot and the ground). This is a crucial block, as any errors at this stage would propagate to the subsequent blocks. The selected algorithm consists of a succession of several signal smoothing and enhancement stages of the raw signal. Locomotion detection is based on the detection of peaks associated with heel-strike events. To be insensitive to sensor placement and orientation, the algorithm is based on the acceleration norm ($accN$), which includes all the components of the acceleration (i.e., vertical acc_v , antero-posterior acc_{AP} and medio-lateral acc_{ML} acceleration), defined as:

$$accN = \sqrt{acc_v^2 + acc_{AP}^2 + acc_{ML}^2}.$$

The raw acceleration norm is first resampled at 40 Hz to correspond to a lower frequency adapted for long-term monitoring setups; the signal is then detrended and low-pass filtered [310]. The $accN$ is further processed with a convolutional wavelet transform (scale 10, gauss2) and three consecutive mild Gaussian weighted moving filters over a window length corresponding to about 0.25 s, to obtain $accN$ enhanced ($accN-enh$). A data-adaptive threshold is used to identify the potential step-related

peaks, by means of the Hilbert transform and adaptive smoothed envelope of the *accN-enh*, followed by detection and statistical description of peak amplitudes in *accN-enh* during all detected walking periods. The threshold is selected as the 5th percentile of peak amplitude distribution (i.e., the value over which 95% of the peaks may be found) [311].

Stride detection

The algorithm selected for the step/stride detection combines different processing techniques. The vertical component of the acceleration is firstly detrended and low-pass filtered (cut-off frequency $f_c=3.2$ Hz), then numerically integrated and differentiated using a continuous wavelet transform (scale 9, gauss2) [312]. Each step is identified by detecting the maximal peaks between zero-crossings, that represent the initial contacts [313,314].

Stride length

The stride length estimation algorithm is a biomechanical-based approach. The inverted pendulum model is used to describe the human body segments (and their biomechanics) during walking. Assuming a compass gait type, the centre of mass (CoM) changes in height and moves in the sagittal plane following a circular trajectory. The step length depends on this height changes according to the following equation: $SL = 2\sqrt{2lh - h^2}$, where l is the pendulum length (i.e., the leg length) and h is equal to the change in the height of the CoM. The vertical component of the acceleration is double integrated and high-pass filtered with a fourth-order Butterworth (respectively, the acceleration is filtered using the cut-off frequency $f_c = 0.1$ Hz, and the vertical speed using the cut-off frequency $f_c=1$ Hz) to extract the vertical displacement of the CoM. The h parameter is then computed as the difference between the highest and the lowest position of the CoM during a step cycle [313–315].

Cadence

To extract cadence, a time-domain approach is employed, which includes the following list of steps:

- Savitzky-Golay filter (order=7, frame length =21)
- Detrending, LPF (FIR, $f_c=3.2$ Hz)
- Continuous Wavelet Transform (scale 10, 'gaus2')
- Savitzky-Golay filter (order=5, frame length =11)
- Continuous Wavelet Transform (scale 10, 'gaus2')
- Gaussian smoothing
- Morphological filters: opening and closing

- Difference signal: closing-opening

Savitzky-Golay filter is used for smoothing noise data by fitting successive adjacent data points with a low-degree polynomial by the method of the linear least squares (i.e., the convolution process).

Continuous Wavelet Transform is a frequency domain filtering of the signal that allows to achieve a varying time-frequency localization. It uses short wavelets to ensure high-time localization at the expense of limited information on the frequency content, and conversely, long wavelets to analyse low frequencies precisely but with limited time localization.

Opening and closing are methods based on the morphological filters and are defined as follows:

$$\text{Opening: } (f \circ k)(m) = [(f \ominus k) \oplus k](m)$$

$$\text{Closing: } (f \cdot k)(m) = [(f \oplus k) \ominus k](m),$$

where $f(m)$ is the measured data and $k(m)$ is a structuring element [309].

Real-World walking speed

Real-World walking speed is mathematically computed by combining the cadence and the stride length.

$$\text{WalkingSpeed}_i = \frac{\sum_{k=1}^{n_STRIDE_i} \text{Stride_Speed}_k}{n_STRIDE_i},$$

Where $i=1, \dots, n$ are the different WBs, Stride_speed_k is the stride speed of the k – stride in the relevant i – WB, n_STRIDE_i is the number of correct strides identified in the relevant i – WB.

Left/Right stride detection

Initial contacts are classified left or right using a pre-trained linear support vector machine with features from vertical and anterior-posterior angular velocity signals.

A combination of the vertical, gyr_v , and the anterior-posterior, gyr_{ap} , angular acceleration (i.e., a subtraction, $gyr_{comb} = gyr_v - gyr_{ap}$), was used as input for the left/right detection. A support-vector machine approach in an individual leave-one-participant-out cross-validation was implemented and evaluated.

Table A.1 summarises all the main information about each block (e.g., input).

Table A.1 Summary of the main information about the algorithms used to extract each DMO.

Modules	Input	Output	Data-driven or mechanistic	Refs
---------	-------	--------	----------------------------------	------

		vertical acceleration	DD	[311]
Gait sequence detection	vertical acceleration	anterior-posterior acceleration		
	anterior-posterior acceleration	window size for convolution		
	medio-lateral acceleration	overlap of windows		
		motion threshold		
Step detection	Vertical acceleration	detected N initial contacts time relative to the beginning of the recording or the test/trial	M	[310–312]
	array of gait sequences	refinement of the beginning of the gait sequences provided from the Gait sequence detection step refinement of the termination of the gait sequences provided from the Gait sequence detection step		
Step length		A vector structure containing: start of N-th GS stop of N-th GS	M	[313–315]
	Vertical orientation	estimated stride length-per-second [m] of N-th GS		
	sampling frequency	mean of stride length [m] estimated for N-th GS		
	array of gait sequences	std of stride length [m] estimated for N-th GS		
	initial contacts time	covered distance [m] of N-th GS		
	leg length	shows if there is a warning for N-th GS (1 means yes)		
	calibration factors			
Cadence		A vector structure containing: start of N-th GS stop of N-th GS	M	[309,311]
	norm of the 3d-accelerometer signal	estimated cadence-per-sec [steps/min] of N-th GS		
	sampling frequency	mean of cadence [steps/min] estimated for N-th GS		
	array of gait sequences	std of cadence [steps/min] estimated for N-th GS		
		number of steps of N-th GS		
Real-world walking speed	cadence array of gait sequences	Real-world walking speed	M	
Left/right stride detection	vertical angular velocity anterior-posterior angular velocity	Detected initial contacts time with the “L” and “R” labels assigned	DD	[316]

array of initial contacts

References

1. Webber SC, Porter MM, Menec VH. Mobility in older adults: a comprehensive framework. *Gerontologist*. 2010;50: 443–450. doi:10.1093/geront/gnq013
2. Richardson J, Letts L, Chan D, Officer A, Wojkowski S, Oliver D, et al. Monitoring physical functioning as the sixth vital sign: evaluating patient and practice engagement in chronic illness care in a primary care setting--a quasi-experimental design. *BMC Family Practice*. 2012;13: 29. doi:10.1186/1471-2296-13-29
3. Hoffman JM, Ciol MA, Huynh M, Chan L. Estimating Transition Probabilities in Mobility and Total Costs for Medicare Beneficiaries. *Archives of Physical Medicine and Rehabilitation*. 2010;91: 1849–1855. doi:10.1016/j.apmr.2010.08.010
4. Metz DH. Mobility of older people and their quality of life. *Transport Policy*. 2000;7: 149–152. doi:10.1016/S0967-070X(00)00004-4
5. Yeom HA, Fleury J, Keller C. Risk Factors for Mobility Limitation in Community-Dwelling Older Adults: A Social Ecological Perspective. *Geriatric Nursing*. 2008;29: 133–140. doi:10.1016/j.gerinurse.2007.07.002
6. WHO. Ageing and health. 2022. Available: <https://www.who.int/news-room/fact-sheets/detail/ageing-and-health>
7. Stenholm S, Shardell M, Bandinelli S, Guralnik JM, Ferrucci L. Physiological Factors Contributing to Mobility Loss Over 9 Years of Follow-Up—Results From the InCHIANTI Study. *The Journals of Gerontology: Series A*. 2015;70. doi:10.1093/gerona/glv004
8. Sorond FA, Cruz-Almeida Y, Clark DJ, Viswanathan A, Scherzer CR, De Jager P, et al. Aging, the Central Nervous System, and Mobility in Older Adults: Neural Mechanisms of Mobility Impairment. *The Journals of Gerontology: Series A*. 2015;70: 1526–1532. doi:10.1093/gerona/glv130
9. Triggs LN, Rogers J. *The Musculoskeletal System and Human Movement*. Orthopaedic and trauma nursing. John Wiley & Sons, Ltd; 2014. pp. 27–47. doi:10.1002/9781118941263.ch4
10. Wang Z, Cui K, Song R, Li X, Qi X, Buchman AS, et al. Influence of Cardiovascular Risk Burden on Motor Function Among Older Adults: Mediating Role of Cardiovascular Diseases Accumulation and Cognitive Decline. *Frontiers in Medicine*. 2022;9. Available: <https://www.frontiersin.org/articles/10.3389/fmed.2022.856260>
11. Abe T, Suzuki T, Yoshida H, Shimada H, Inoue N. The Relationship Between Pulmonary Function and Physical Function and Mobility in Community-Dwelling Elderly Women Aged 75 Years or Older. *Journal of Physical Therapy Science*. 2011;23: 443–449. doi:10.1589/jpts.23.443
12. Dillon CF, Gu Q, Hoffman HJ, Ko C-W. Vision, hearing, balance, and sensory impairment in Americans aged 70 years and over: United States, 1999-2006. *NCHS Data Brief*. 2010.
13. Iolascon G, Paoletta M, Liguori S, Curci C, Moretti A. Neuromuscular Diseases and Bone. *Front Endocrinol (Lausanne)*. 2019;10: 794. doi:10.3389/fendo.2019.00794
14. Baradaran N, Tan SN, Liu A, Ashoori A, Palmer SJ, Wang ZJ, et al. Parkinson's Disease Rigidity: Relation to Brain Connectivity and Motor Performance. *Front Neurol*. 2013;4: 67. doi:10.3389/fneur.2013.00067
15. Fong SSM, Chung LMY, Bae Y-H, Vackova D, Ma AWW, Liu KPY. Neuromuscular Processes in the Control of Posture in Children with Developmental Coordination Disorder: Current Evidence and Future Research Directions. *Curr Dev Disord Rep*. 2018;5: 43–48. doi:10.1007/s40474-018-

16. Kerr C, Bottomley C, Shingler S, Giangregorio L, de Freitas HM, Patel C, et al. The importance of physical function to people with osteoporosis. *Osteoporos Int.* 2017;28: 1597–1607. doi:10.1007/s00198-017-3911-9
17. Selected Health Conditions and Likelihood of Improvement with Treatment. Washington, D.C.: National Academies Press; 2020. doi:10.17226/25662
18. Visser M, Goodpaster BH, Kritchevsky SB, Newman AB, Nevitt M, Rubin SM, et al. Muscle mass, muscle strength, and muscle fat infiltration as predictors of incident mobility limitations in well-functioning older persons. *J Gerontol A Biol Sci Med Sci.* 2005;60: 324–333. doi:10.1093/gerona/60.3.324
19. Clark BC, Manini TM. Functional consequences of sarcopenia and dynapenia in the elderly. *Curr Opin Clin Nutr Metab Care.* 2010;13: 271–276. doi:10.1097/MCO.0b013e328337819e
20. Clark BC, Manini TM. What is dynapenia? *Nutrition.* 2012;28: 495–503. doi:10.1016/j.nut.2011.12.002
21. Lauretani F, Russo CR, Bandinelli S, Bartali B, Cavazzini C, Di Iorio A, et al. Age-associated changes in skeletal muscles and their effect on mobility: an operational diagnosis of sarcopenia. *Journal of Applied Physiology.* 2003;95: 1851–1860. doi:10.1152/jappphysiol.00246.2003
22. Fried LP, Guralnik JM. Disability in Older Adults: Evidence Regarding Significance, Etiology, and Risk. *Journal of the American Geriatrics Society.* 1997;45: 92–100. doi:10.1111/j.1532-5415.1997.tb00986.x
23. van den Berg-Emons H (Rita), Bussmann J (Hans), Balk A, Keijzer-Oster D, Stam H. Level of Activities Associated With Mobility During Everyday Life in Patients With Chronic Congestive Heart Failure as Measured With an “Activity Monitor.” *Physical Therapy.* 2001;81: 1502–1511. doi:10.1093/ptj/81.9.1502
24. Chen P-C, Yang T-H, Wu P-J, Wang L-Y, Chen S-M. Mobility Status Plays an Important Role in the Risk of Cardiovascular Rehospitalizations in Patients with Heart Failure Undergoing Cardiac Rehabilitation: A Retrospective Cohort Study. *Journal of Personalized Medicine.* 2022;12: 675. doi:10.3390/jpm12050675
25. Noale M, Veronese N, Smith L, Ungar A, Fumagalli S, Maggi S. Associations between cardiac arrhythmia, incident disability in activities of daily living and physical performance: the ILSA study. *J Geriatr Cardiol.* 2020;17: 127–132. doi:10.11909/j.issn.1671-5411.2020.03.008
26. Crişan AF, Oancea C, Timar B, Fira-Mladinescu O, Tudorache V. Balance Impairment in Patients with COPD. *PLOS ONE.* 2015;10: e0120573. doi:10.1371/journal.pone.0120573
27. Maltais F, LeBlanc P, Jobin J, Casaburi R. Peripheral Muscle Dysfunction in Chronic Obstructive Pulmonary Disease. *Clinics in Chest Medicine.* 2000;21: 665–677. doi:10.1016/S0272-5231(05)70176-3
28. Fisher DE, Ward MM, Hoffman HJ, Li C-M, Cotch MF. Impact of Sensory Impairments on Functional Disability in Adults With Arthritis. *American Journal of Preventive Medicine.* 2016;50: 454–462. doi:10.1016/j.amepre.2015.07.014
29. Swenor BK, Muñoz B, West SK. A longitudinal study of the association between visual impairment and mobility performance in older adults: the salisbury eye evaluation study. *Am J Epidemiol.* 2014;179: 313–322. doi:10.1093/aje/kwt257
30. Chen X, Mao G, Leng SX. Frailty syndrome: an overview. *Clin Interv Aging.* 2014;9: 433–441.

doi:10.2147/CIA.S45300

31. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, et al. Frailty in Older Adults: Evidence for a Phenotype. *The Journals of Gerontology: Series A*. 2001;56: M146–M157. doi:10.1093/gerona/56.3.M146
32. Frailty in Older People. *Lancet*. 2013;381: 752–762. doi:10.1016/S0140-6736(12)62167-9
33. Bergland A, Jørgensen L, Emaus N, Strand BH. Mobility as a predictor of all-cause mortality in older men and women: 11.8 year follow-up in the Tromsø study. *BMC Health Services Research*. 2017;17. doi:10.1186/s12913-016-1950-0
34. Tolea MI, Galvin JE. The relationship between mobility dysfunction staging and global cognitive performance. *Alzheimer Disease and Associated Disorders*. 2016;30: 230–236. doi:10.1097/WAD.0000000000000136
35. Mahmoudi R, Novella J-L, Manckoundia P, Ahssaini F, Lang P-O, Blanchard F, et al. Is functional mobility an independent mortality risk factor in subjects with dementia? *Maturitas*. 2017;103: 65–70. doi:10.1016/j.maturitas.2017.06.031
36. Wallace LMK, Theou O, Godin J, Andrew MK, Bennett DA, Rockwood K. Investigation of frailty as a moderator of the relationship between neuropathology and dementia in Alzheimer’s disease: a cross-sectional analysis of data from the Rush Memory and Aging Project. *The Lancet Neurology*. 2019;18: 177–184. doi:10.1016/S1474-4422(18)30371-5
37. Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, et al. Gait Speed and Survival in Older Adults. *JAMA*. 2011;305: 50–58. doi:10.1001/jama.2010.1923
38. Perera S, Patel KV, Rosano C, Rubin SM, Satterfield S, Harris T, et al. Gait Speed Predicts Incident Disability: A Pooled Analysis. *The Journals of Gerontology: Series A*. 2016;71: 63–71. doi:10.1093/gerona/glv126
39. Rochester L, Mazzà C, Mueller A, Caulfield B, McCarthy M, Becker C, et al. A Roadmap to Inform Development, Validation and Approval of Digital Mobility Outcomes: The Mobilise-D Approach. *Digital Biomarkers*. 2020;4: 13–27. doi:10.1159/000512513
40. Chiari L, Croce UD, Leardini A, Cappozzo A. Human movement analysis using stereophotogrammetry: Part 2: Instrumental errors. *Gait & Posture*. 2005;21: 197–211. doi:10.1016/j.gaitpost.2004.04.004
41. Cappozzo A. Instrumental observation of human movement: historical development. *Three-Dimensional Analysis of Human Locomotion*. 1998; 1–25.
42. Ancillao A. Analysis and Measurement of Human Motion: Modern Protocols and Clinical Considerations. *Journal of Robotics and Mechanical Engineering Research*. 2016;1: 30–37. doi:10.24218/jrmer.2016.19
43. Cappozzo A, Della Croce U, Leardini A, Chiari L. Human movement analysis using stereophotogrammetry. Part 1: theoretical background. *Gait Posture*. 2005;21: 186–196. doi:10.1016/j.gaitpost.2004.01.010
44. Medved V. *Measurement of Human Locomotion*. CRC Press; 2000.
45. Cappozzo A. Gait analysis methodology. *Human Movement Science*. 1984;3: 27–50. doi:10.1016/0167-9457(84)90004-6
46. Di Marco R, Rossi S, Castelli E, Patanè F, Mazzà C, Cappa P. Effects of the calibration procedure on the metrological performances of stereophotogrammetric systems for human movement analysis. *Measurement*. 2017;101: 265–271. doi:10.1016/j.measurement.2016.01.008

47. Capozzo A, Berme N. *Biomechanics of Human Movement: Applications in Rehabilitation, Sports and Ergonomics*. Bertec Corporation; 1990.
48. Basmajian JV. Muscles Alive—their functions revealed by electromyography. *Postgrad Med J*. 1963;39: 162.
49. Hubble RP, Silburn PA, Naughton GA, Cole MH. Assessing stability in mild and moderate Parkinson's disease: Can clinical measures provide insight? *Gait & Posture*. 2016;49: 7–13. doi:10.1016/j.gaitpost.2016.06.002
50. Podsiadlo D, Richardson S. The Timed “Up & Go”: A Test of Basic Functional Mobility for Frail Elderly Persons. *Journal of the American Geriatrics Society*. 1991;39: 142–148. doi:10.1111/j.1532-5415.1991.tb01616.x
51. Deathe AB, Miller WC. The L test of functional mobility: measurement properties of a modified version of the timed “up & go” test designed for people with lower-limb amputations. *Physical therapy*. 2005;85. Available: <https://pubmed.ncbi.nlm.nih.gov/15982169/>
52. Maetzler W, Rochester L, Bhidayasiri R, Espay AJ, Sánchez-Ferro A, van Uem JMT. Modernizing Daily Function Assessment in Parkinson's Disease Using Capacity, Perception, and Performance Measures. *Movement Disorders*. 2021;36: 76–82. doi:10.1002/mds.28377
53. Warmerdam E, Hausdorff JM, Atrsaei A, Zhou Y, Mirelman A, Aminian K, et al. Long-term unsupervised mobility assessment in movement disorders. *The Lancet Neurology*. 2020;19: 462–470. doi:10.1016/S1474-4422(19)30397-7
54. Hillel I, Gazit E, Nieuwboer A, Avanzino L, Rochester L, Cereatti A, et al. Is every-day walking in older adults more analogous to dual-task walking or to usual walking? Elucidating the gaps between gait performance in the lab and during 24/7 monitoring. *Eur Rev Aging Phys Act*. 2019;16: 6. doi:10.1186/s11556-019-0214-5
55. Takayanagi N, Sudo M, Yamashiro Y, Lee S, Kobayashi Y, Niki Y, et al. Relationship between Daily and In-laboratory Gait Speed among Healthy Community-dwelling Older Adults. *Sci Rep*. 2019;9: 3496. doi:10.1038/s41598-019-39695-0
56. Pareés I, Saifee TA, Kassavetis P, Kojovic M, Rubio-Agusti I, Rothwell JC, et al. Believing is perceiving: mismatch between self-report and actigraphy in psychogenic tremor. *Brain*. 2012;135: 117–123. doi:10.1093/brain/awr292
57. Jones SMW, Shulman LJ, Richards JE, Ludman EJ. Mechanisms for the testing effect on patient-reported outcomes. *Contemp Clin Trials Commun*. 2020;18: 100554. doi:10.1016/j.conctc.2020.100554
58. van Gilst MM, Cramer IC, Bloem BR, Overeem S, Faber MJ. A grounded theory study on the influence of sleep on Parkinson's symptoms. *BMC Res Notes*. 2016;9: 299. doi:10.1186/s13104-016-2114-3
59. Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Movement Disorders*. 2008;23: 2129–2170. doi:10.1002/mds.22340
60. Hobart JC, Riazi A, Lamping DL, Fitzpatrick R, Thompson AJ. Measuring the impact of MS on walking ability: The 12-Item MS Walking Scale (MSWS-12). *Neurology*. 2003;60: 31–36. doi:10.1212/WNL.60.1.31
61. Taraldsen K, Vereijken B, Thingstad P, Sletvold O, Helbostad JL. Multiple Days of Monitoring Are Needed to Obtain a Reliable Estimate of Physical Activity in Hip-Fracture Patients. *Journal of Aging*

and Physical Activity. 2014;22: 173–177. doi:10.1123/japa.2012-0130

62. Ercan İP, TiMur Ş. Changing Terminology of Definition and Design of Wearable Technology Products. 2020;8.
63. Thorp EO. *Beat the Dealer: A Winning Strategy for the Game of Twenty-One*. Knopf Doubleday Publishing Group; 1966.
64. Thorp EO. The invention of the first wearable computer. *Digest of Papers Second International Symposium on Wearable Computers (Cat No98EX215)*. Pittsburgh, PA, USA: IEEE Comput. Soc; 1998. pp. 4–8. doi:10.1109/ISWC.1998.729523
65. Ometov A, Shubina V, Klus L, Skibińska J, Saafi S, Pascacio P, et al. A Survey on Wearable Technology: History, State-of-the-Art and Current Challenges. *Computer Networks*. 2021;193: 108074. doi:10.1016/j.comnet.2021.108074
66. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity*. 2015;12: 159. doi:10.1186/s12966-015-0314-1
67. Brickwood K-J, Watson G, O'Brien J, Williams AD. Consumer-Based Wearable Activity Trackers Increase Physical Activity Participation: Systematic Review and Meta-Analysis. *JMIR mHealth and uHealth*. 2019;7: e11819. doi:10.2196/11819
68. Washington WD, Banna KM, Gibson AL. Preliminary efficacy of prize-based contingency management to increase activity levels in healthy adults. *Journal of Applied Behavior Analysis*. 2014;47: 231–245. doi:10.1002/jaba.119
69. El-Amrawy F, Nounou MI. Are Currently Available Wearable Devices for Activity Tracking and Heart Rate Monitoring Accurate, Precise, and Medically Beneficial? *Healthc Inform Res*. 2015;21: 315–320. doi:10.4258/hir.2015.21.4.315
70. Kuzubasoglu BA, Sayar E, Cochrane C, Koncar V, Bahadir SK. Wearable temperature sensor for human body temperature detection. *J Mater Sci: Mater Electron*. 2021;32: 4784–4797. doi:10.1007/s10854-020-05217-2
71. Coughlin SS, Stewart J. Use of Consumer Wearable Devices to Promote Physical Activity: A Review of Health Intervention Studies. *J Environ Health Sci*. 2016;2: 10.15436/2378-6841.16.1123.
72. Storm FA, Heller BW, Mazzà C. Step detection and activity recognition accuracy of seven physical activity monitors. *PLoS One*. 2015;10: e0118723. doi:10.1371/journal.pone.0118723
73. Bonato P. Wearable sensors/systems and their impact on biomedical engineering. *IEEE Engineering in Medicine and Biology Magazine*. 2003;22: 18–20. doi:10.1109/MEMB.2003.1213622
74. Sciacqua A, Valentini M, Gualtieri A, Perticone F, Faini A, Zacharioudakis G, et al. Validation of a flexible and innovative platform for the home monitoring of heart failure patients: Preliminary results. 2009 36th Annual Computers in Cardiology Conference (CinC). 2009. pp. 97–100. Available: <https://ieeexplore.ieee.org/document/5445460>
75. Sung M, Marci C, Pentland A. Wearable feedback systems for rehabilitation. *Journal of NeuroEngineering and Rehabilitation*. 2005;2: 17. doi:10.1186/1743-0003-2-17
76. Warmerdam E, Romijnders R, Geritz J, Elshehabi M, Maetzler C, Otto JC, et al. 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*. 2021;21: 5833. doi:10.3390/s21175833
77. Mazzà C, Alcock L, Aminian K, Becker C, Bertuletti S, Bonci T, et al. Technical validation of real-

- world monitoring of gait: a multicentric observational study. *BMJ Open*. 2021;11: e050785. doi:10.1136/bmjopen-2021-050785
78. O'Day J, Lee M, Seagers K, Hoffman S, Jih-Schiff A, Kidziński Ł, et al. Assessing inertial measurement unit locations for freezing of gait detection and patient preference. *Journal of NeuroEngineering and Rehabilitation*. 2022;19. doi:10.1186/s12984-022-00992-x
 79. Storm FA, Buckley CJ, Mazzà C. Gait event detection in laboratory and real life settings: Accuracy of ankle and waist sensor based methods. *Gait & Posture*. 2016;50: 42–46. doi:10.1016/j.gaitpost.2016.08.012
 80. Puri A, Kim B, Nguyen O, Stolee P, Tung J, Lee J. User Acceptance of Wrist-Worn Activity Trackers Among Community-Dwelling Older Adults: Mixed Method Study. *JMIR mHealth and uHealth*. 2017;5: e8211. doi:10.2196/mhealth.8211
 81. Del Din S, Godfrey A, Rochester L. 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*. 2016;20: 838–847. doi:10.1109/JBHI.2015.2419317
 82. Bonci T, Keogh A, Del Din S, Scott K, Mazzà C, on behalf of the Mobilise-D consortium. An Objective Methodology for the Selection of a Device for Continuous Mobility Assessment. *Sensors*. 2020;20: 6509. doi:10.3390/s20226509
 83. Hickey A, Gunn E, Alcock L, Del Din S, Godfrey A, Rochester L, et al. Validity of a wearable accelerometer to quantify gait in spinocerebellar ataxia type 6. *Physiol Meas*. 2016;37: N105–N117. doi:10.1088/0967-3334/37/11/N105
 84. Dijkstra B, Zijlstra W, Scherder E, Kamsma Y. 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*. 2008;37: 436–441. doi:10.1093/ageing/afn097
 85. Furlanetto KC, Bisca GW, Oldemberg N, Sant'Anna TJ, Morakami FK, Camillo CA, et al. Step Counting and Energy Expenditure Estimation in Patients With Chronic Obstructive Pulmonary Disease and Healthy Elderly: Accuracy of 2 Motion Sensors. *Archives of Physical Medicine and Rehabilitation*. 2010;91: 261–267. doi:10.1016/j.apmr.2009.10.024
 86. Angelini L, Hodgkinson W, Smith C, Dodd JM, Sharrack B, Mazzà C, et al. Wearable sensors can reliably quantify gait alterations associated with disability in people with progressive multiple sclerosis in a clinical setting. *J Neurol*. 2020;267: 2897–2909. doi:10.1007/s00415-020-09928-8
 87. Angelini L, Buckley E, Bonci T, Radford A, Sharrack B, Paling D, et al. A Multifactorial Model of Multiple Sclerosis Gait and Its Changes Across Different Disability Levels. *IEEE Transactions on Biomedical Engineering*. 2021;68: 3196–3204. doi:10.1109/TBME.2021.3061998
 88. Soltani A, Aminian K, Mazza C, Cereatti A, Palmerini L, Bonci T, et al. Algorithms for Walking Speed Estimation Using a Lower-Back-Worn Inertial Sensor: A Cross-Validation on Speed Ranges. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2021;29: 1955–1964. doi:10.1109/TNSRE.2021.3111681
 89. Arnold EM, Ward SR, Lieber RL, Delp SL. A Model of the Lower Limb for Analysis of Human Movement. *Ann Biomed Eng*. 2010;38: 269–279. doi:10.1007/s10439-009-9852-5
 90. Delp SL, Loan JP, Hoy MG, Zajac FE, Topp EL, Rosen JM. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. *IEEE Transactions on Biomedical Engineering*. 1990;37: 757–767. doi:10.1109/10.102791
 91. Hamner SR, Seth A, Delp SL. Muscle contributions to propulsion and support during running.

Journal of Biomechanics. 2010;43: 2709–2716. doi:10.1016/j.jbiomech.2010.06.025

92. Rajagopal A, Dembia CL, DeMers MS, Delp DD, Hicks JL, Delp SL. Full-Body Musculoskeletal Model for Muscle-Driven Simulation of Human Gait. *IEEE Trans Biomed Eng.* 2016;63: 2068–2079. doi:10.1109/TBME.2016.2586891
93. Ward SR, Eng CM, Smallwood LH, Lieber RL. Are Current Measurements of Lower Extremity Muscle Architecture Accurate? *Clin Orthop Relat Res.* 2009;467: 1074–1082. doi:10.1007/s11999-008-0594-8
94. Carbone V, Fluit R, Pellikaan P, van der Krogt MM, Janssen D, Damsgaard M, et al. TLEM 2.0 - a comprehensive musculoskeletal geometry dataset for subject-specific modeling of lower extremity. *J Biomech.* 2015;48: 734–741. doi:10.1016/j.jbiomech.2014.12.034
95. Brand RA, Pedersen DR, Davy DT, Kotzar GM, Heiple KG, Goldberg VM. Comparison of hip force calculations and measurements in the same patient. *J Arthroplasty.* 1994;9: 45–51. doi:10.1016/0883-5403(94)90136-8
96. Garner BA, Pandy MG. Musculoskeletal model of the upper limb based on the visible human male dataset. *Comput Methods Biomech Biomed Engin.* 2001;4: 93–126. doi:10.1080/10255840008908000
97. Gu L, Xu J, Peters TM. Novel multistage three-dimensional medical image segmentation: methodology and validation. *IEEE Trans Inf Technol Biomed.* 2006;10: 740–748. doi:10.1109/titb.2006.875665
98. Kainz H, Goudriaan M, Falisse A, Huenaerts C, Desloovere K, De Groote F, et al. The influence of maximum isometric muscle force scaling on estimated muscle forces from musculoskeletal models of children with cerebral palsy. *Gait & Posture.* 2018;65: 213–220. doi:10.1016/j.gaitpost.2018.07.172
99. Correa TA, Baker R, Graham HK, Pandy MG. Accuracy of generic musculoskeletal models in predicting the functional roles of muscles in human gait. *J Biomech.* 2011;44: 2096–2105. doi:10.1016/j.jbiomech.2011.05.023
100. Scheys L, Van Campenhout A, Spaepen A, Suetens P, Jonkers I. Personalized MR-based musculoskeletal models compared to rescaled generic models in the presence of increased femoral anteversion: Effect on hip moment arm lengths. *Gait & Posture.* 2008;28: 358–365. doi:10.1016/j.gaitpost.2008.05.002
101. Valente G, Crimi G, Vanella N, Schileo E, Taddei F. nmsBuilder: Freeware to create subject-specific musculoskeletal models for OpenSim. *Computer Methods and Programs in Biomedicine.* 2017;152: 85–92. doi:10.1016/j.cmpb.2017.09.012
102. Modenese L, Renault J-B. Automatic generation of personalised skeletal models of the lower limb from three-dimensional bone geometries. *Journal of Biomechanics.* 2021;116: 110186. doi:10.1016/j.jbiomech.2020.110186
103. Delp SL, Anderson FC, Arnold AS, Loan P, Habib A, John CT, et al. OpenSim: Open-Source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Transactions on Biomedical Engineering.* 2007;54: 1940–1950. doi:10.1109/TBME.2007.901024
104. Damsgaard M, Rasmussen J, Christensen ST, Surma E, de Zee M. Analysis of musculoskeletal systems in the AnyBody Modeling System. *Simulation Modelling Practice and Theory.* 2006;14: 1100–1111. doi:10.1016/j.simpat.2006.09.001
105. Modenese L, Montefiori E, Wang A, Wesarg S, Viceconti M, Mazzà C. Investigation of the dependence of joint contact forces on musculotendon parameters using a codified workflow for

- image-based modelling. *Journal of Biomechanics*. 2018;73: 108–118. doi:10.1016/j.jbiomech.2018.03.039
106. Rüdiger HA, Guillemin M, Latypova A, Terrier A. Effect of changes of femoral offset on abductor and joint reaction forces in total hip arthroplasty. *Arch Orthop Trauma Surg*. 2017;137: 1579–1585. doi:10.1007/s00402-017-2788-6
 107. Nettrour JF, Razu SS, Keeney JA, Guess TM. Femoral Component Malrotation Produces Quadriceps Weakness and Impaired Ambulatory Function following Total Knee Arthroplasty: Results of a Forward-Dynamic Computer Model. *J Knee Surg*. 2019. doi:10.1055/s-0039-1692486
 108. Pitto L, Kainz H, Falisse A, Wesseling M, Van Rossom S, Hoang H, et al. SimCP: A Simulation Platform to Predict Gait Performance Following Orthopedic Intervention in Children With Cerebral Palsy. *Front Neurobot*. 2019;13: 54. doi:10.3389/fnbot.2019.00054
 109. Martelli S, Calvetti D, Somersalo E, Viceconti M. Stochastic modelling of muscle recruitment during activity. *Interface Focus*. 2015;5: 20140094. doi:10.1098/rsfs.2014.0094
 110. Montefiori E, Modenese L, Di Marco R, Magni-Manzoni S, Malattia C, Petrarca M, et al. An image-based kinematic model of the tibiotalar and subtalar joints and its application to gait analysis in children with Juvenile Idiopathic Arthritis. *Journal of Biomechanics*. 2019;85: 27–36. doi:10.1016/j.jbiomech.2018.12.041
 111. Montefiori E, Modenese L, Di Marco R, Magni-Manzoni S, Malattia C, Petrarca M, et al. Linking Joint Impairment and Gait Biomechanics in Patients with Juvenile Idiopathic Arthritis. *Ann Biomed Eng*. 2019;47: 2155–2167. doi:10.1007/s10439-019-02287-0
 112. Ong CF, Geijtenbeek T, Hicks JL, Delp SL. Predicting gait adaptations due to ankle plantarflexor muscle weakness and contracture using physics-based musculoskeletal simulations. *PLOS Computational Biology*. 2019;15: e1006993. doi:10.1371/journal.pcbi.1006993
 113. Wesseling M, de Groote F, Meyer C, Corten K, Simon J-P, Desloovere K, et al. Gait alterations to effectively reduce hip contact forces. *Journal of Orthopaedic Research*. 2015;33: 1094–1102. doi:10.1002/jor.22852
 114. Knarr BA, Reisman DS, Binder-Macleod SA, Higginson JS. Understanding compensatory strategies for muscle weakness during gait by simulating activation deficits seen post-stroke. *Gait Posture*. 2013;38: 270–275. doi:10.1016/j.gaitpost.2012.11.027
 115. Viceconti M, Juárez MA, Curreli C, Pennisi M, Russo G, Pappalardo F. Credibility of In Silico Trial Technologies—A Theoretical Framing. *IEEE Journal of Biomedical and Health Informatics*. 2020;24: 4–13. doi:10.1109/JBHI.2019.2949888
 116. Pappalardo F, Russo G, Tshinanu FM, Viceconti M. In silico clinical trials: concepts and early adoptions. *Brief Bioinform*. 2019;20: 1699–1708. doi:10.1093/bib/bby043
 117. Viceconti M, Cobelli C, Haddad T, Himes A, Kovatchev B, Palmer M. In silico assessment of biomedical products: The conundrum of rare but not so rare events in two case studies. *Proc Inst Mech Eng H*. 2017;231: 455–466. doi:10.1177/0954411917702931
 118. Patterson EA, Whelan MP. A framework to establish credibility of computational models in biology. *Prog Biophys Mol Biol*. 2017;129: 13–19. doi:10.1016/j.pbiomolbio.2016.08.007
 119. Mirams GR, Pathmanathan P, Gray RA, Challenor P, Clayton RH. Uncertainty and variability in computational and mathematical models of cardiac physiology. *The Journal of Physiology*. 2016;594: 6833–6847. doi:10.1113/JP271671
 120. Contrera JF. Validation of Toxtree and SciQSAR in silico predictive software using a publicly

available benchmark mutagenicity database and their applicability for the qualification of impurities in pharmaceuticals. *Regulatory Toxicology and Pharmacology*. 2013;67: 285–293. doi:10.1016/j.yrtph.2013.08.008

121. Hariharan P, D'Souza GA, Horner M, Morrison TM, Malinauskas RA, Myers MR. Use of the FDA nozzle model to illustrate validation techniques in computational fluid dynamics (CFD) simulations. *PLOS ONE*. 2017;12: e0178749. doi:10.1371/journal.pone.0178749
122. European Medicine Agency. Qualification of novel methodologies for drug development: guidance to applicants. 2014.
123. Food & Drug Administration. Qualification Process for Drug Development Tools. Guidance for Industry and FDA Staff.
124. Antich-Isern P, Caro-Barri J, Aparicio-Blanco J. The combination of medical devices and medicinal products revisited from the new European legal framework. *International Journal of Pharmaceutics*. 2021;607: 120992. doi:10.1016/j.ijpharm.2021.120992
125. Council Directive 65/65/EEC of 26 January 1965 on the approximation of provisions laid down by Law, Regulation or Administrative Action relating to proprietary medicinal products. DD I Jan 26, 1965. Available: <http://data.europa.eu/eli/dir/1965/65/oj/eng>
126. Council Directive 93/42/EEC of 14 June 1993 concerning medical devices. OJ L Jun 14, 1993. Available: <http://data.europa.eu/eli/dir/1993/42/oj/eng>
127. Van Norman GA. Drugs and Devices: Comparison of European and U.S. Approval Processes. *JACC: Basic to Translational Science*. 2016;1: 399–412. doi:10.1016/j.jacbts.2016.06.003
128. Kaladji A, Lucas A, Kervio G, Haigron P, Cardon A. Sizing for Endovascular Aneurysm Repair: Clinical Evaluation of a New Automated Three-Dimensional Software. *Annals of Vascular Surgery*. 2010;24: 912–920. doi:10.1016/j.avsg.2010.03.018
129. Taylor CA, Fonte TA, Min JK. Computational Fluid Dynamics Applied to Cardiac Computed Tomography for Noninvasive Quantification of Fractional Flow Reserve: Scientific Basis. *Journal of the American College of Cardiology*. 2013;61: 2233–2241. doi:10.1016/j.jacc.2012.11.083
130. HeartFlow (2019). CT-Flow. 2019. Available: <https://www.heartflow.com/ct-flow/>
131. Viceconti M, Emili L, Afshari P, Courcelles E, Curreli C, Famaey N, et al. Possible Contexts of Use for In Silico Trials Methodologies: A Consensus-Based Review. *IEEE J Biomed Health Inform*. 2021;25: 3977–3982. doi:10.1109/JBHI.2021.3090469
132. McHugh M, McCaffery F, Casey V. Software process improvement to assist medical device software development organisations to comply with the amendments to the medical device directive. *IET Software*. 2012;6: 431–437. doi:10.1049/iet-sen.2011.0198
133. Zema M, Rosati S, Gioia V, Knaflitz M, Balestra G. Developing medical device software in compliance with regulations. 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2015. pp. 1331–1334. doi:10.1109/EMBC.2015.7318614
134. Pathmanathan P, Gray RA. Validation and Trustworthiness of Multiscale Models of Cardiac Electrophysiology. *Frontiers in Physiology*. 2018;9. Available: <https://www.frontiersin.org/articles/10.3389/fphys.2018.00106>
135. ASME. V&V 40 - 2018: Assessing Credibility of Computational Modeling through Verification and Validation: Application to Medical Devices. 2018.
136. Morrison TM, Hariharan P, Funkhouser CM, Afshari P, Goodin M, Horner M. Assessing Computational Model Credibility Using a Risk-Based Framework: Application to Hemolysis in

- Centrifugal Blood Pumps. *ASAIO Journal*. 2019;65: 349. doi:10.1097/MAT.0000000000000996
137. Mobilise-D. In: Mobilise-D [Internet]. Available: <https://mobilise-d.eu/>
 138. Yang L, Lu K, Forsman M, Lindecrantz K, Seoane F, Ekblom Ö, et al. Evaluation of physiological workload assessment methods using heart rate and accelerometry for a smart wearable system. *Ergonomics*. 2019;62: 694–705. doi:10.1080/00140139.2019.1566579
 139. Ekelund U, Tarp J, Steene-Johannessen J, Hansen BH, Jefferis B, Fagerland MW, et al. Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis. *BMJ*. 2019;366: 14570. doi:10.1136/bmj.14570
 140. Jankovic J. Parkinson's disease: clinical features and diagnosis. *J Neurol Neurosurg Psychiatry*. 2008;79: 368–376. doi:10.1136/jnnp.2007.131045
 141. Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord*. 2015;30: 1591–1601. doi:10.1002/mds.26424
 142. Pringsheim T, Jette N, Frolkis A, Steeves TDL. The prevalence of Parkinson's disease: a systematic review and meta-analysis. *Mov Disord*. 2014;29: 1583–1590. doi:10.1002/mds.25945
 143. Hoehn MM, Yahr MD. Parkinsonism: Onset, progression, and mortality. *Neurology*. 1967;17: 427–427. doi:10.1212/WNL.17.5.427
 144. Goetz CG, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stebbins GT, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan. *Mov Disord*. 2007;22: 41–47. doi:10.1002/mds.21198
 145. Pugliatti M, Rosati G, Carton H, Riise T, Drulovic J, Vécsei L, et al. The epidemiology of multiple sclerosis in Europe. *European Journal of Neurology*. 2006;13: 700–722. doi:10.1111/j.1468-1331.2006.01342.x
 146. Storm FA, Nair KPS, Clarke AJ, Meulen JMV der, Mazzà C. Free-living and laboratory gait characteristics in patients with multiple sclerosis. *PLOS ONE*. 2018;13: e0196463. doi:10.1371/journal.pone.0196463
 147. Delgado-Ortiz L, Polhemus A, Keogh A, Sutton N, Remmele W, Hansen C, et al. Listening to the patients' voice: a conceptual framework of the walking experience. *Age and Ageing*. 2023;52: afac233. doi:10.1093/ageing/afac233
 148. Agustí AGN. Systemic Effects of Chronic Obstructive Pulmonary Disease. *Chronic Obstructive Pulmonary Disease: Pathogenesis to Treatment*. John Wiley & Sons, Ltd; 2000. pp. 242–254. doi:10.1002/0470868678.ch15
 149. Vestbo J, Hurd SS, Agustí AG, Jones PW, Vogelmeier C, Anzueto A, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med*. 2013;187: 347–365. doi:10.1164/rccm.201204-0596PP
 150. European Medicine Agency. Guideline on clinical investigation of medicinal products in the treatment of chronic obstructive pulmonary disease (COPD). 2012.
 151. Gupta N, Pinto LM, Morogan A, Bourbeau J. The COPD assessment test: a systematic review. *European Respiratory Journal*. 2014;44: 873–884. doi:10.1183/09031936.00025214
 152. Meguro M, Barley EA, Spencer S, Jones PW. Development and Validation of an Improved, COPD-Specific Version of the St. George Respiratory Questionnaire. *Chest*. 2007;132: 456–463. doi:10.1378/chest.06-0702

153. Freire AN, Guerra RO, Alvarado B, Guralnik JM, Zunzunegui MV. Validity and reliability of the short physical performance battery in two diverse older adult populations in Quebec and Brazil. *J Aging Health*. 2012;24: 863–878. doi:10.1177/0898264312438551
154. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. *J Gerontol*. 1994;49: M85-94. doi:10.1093/geronj/49.2.m85
155. Parsons N, Griffin XL, Achten J, Costa ML. Outcome assessment after hip fracture: is EQ-5D the answer? *Bone & Joint Research*. 2014;3: 69–75. doi:10.1302/2046-3758.33.2000250
156. Angelini L, Carpinella I, Cattaneo D, Ferrarin M, Gervasoni E, Sharrack B, et al. Is a Wearable Sensor-Based Characterisation of Gait Robust Enough to Overcome Differences Between Measurement Protocols? A Multi-Centric Pragmatic Study in Patients with Multiple Sclerosis. *Sensors*. 2020;20: 79. doi:10.3390/s20010079
157. Vienne A, Barrois RP, Buffat S, Ricard D, Vidal P-P. Inertial Sensors to Assess Gait Quality in Patients with Neurological Disorders: A Systematic Review of Technical and Analytical Challenges. *Front Psychol*. 2017;8: 817. doi:10.3389/fpsyg.2017.00817
158. Vienne-Jumeau A, Quijoux F, Vidal P-P, Ricard D. Wearable inertial sensors provide reliable biomarkers of disease severity in multiple sclerosis: A systematic review and meta-analysis. *Ann Phys Rehabil Med*. 2020;63: 138–147. doi:10.1016/j.rehab.2019.07.004
159. Keogh A, Alcock L, Brown P, Buckley E, Brozgol M, Gazit E, et al. Acceptability of wearable devices for measuring mobility remotely: Observations from the Mobilise-D technical validation study. *Digital Health*. 2023 [cited 2 Jan 2024]. doi:10.1177/20552076221150745
160. Kluge F, Din SD, Cereatti A, Gaßner H, Hansen C, Helbostad JL, et al. Consensus based framework for digital mobility monitoring. *PLOS ONE*. 2021;16: e0256541. doi:10.1371/journal.pone.0256541
161. Mikolaizak AS, Rochester L, Maetzler W, Sharrack B, Demeyer H, Mazzà C, et al. Connecting real-world digital mobility assessment to clinical outcomes for regulatory and clinical endorsement—the Mobilise-D study protocol. *PLoS One*. 2022;17: e0269615. doi:10.1371/journal.pone.0269615
162. Viceconti M, Hernandez Penna S, Dartee W, Mazzà C, Caulfield B, Becker C, et al. Toward a Regulatory Qualification of Real-World Mobility Performance Biomarkers in Parkinson’s Patients Using Digital Mobility Outcomes. *Sensors*. 2020;20: 5920. doi:10.3390/s20205920
163. European Medicine Agency. Qualification opinion on Proactive in chronic obstructive pulmonary disease. 2018. Available: https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/qualification-opinion-proactive-chronic-obstructive-pulmonary-disease-copd_en.pdf
164. European Medicine Agency. Qualification opinion on stride velocity 95th centile as a secondary endpoint in Duchenne Muscular Dystrophy measured by a valid and suitable wearable device. 2019. Available: https://www.ema.europa.eu/en/documents/scientific-guideline/qualification-opinion-stride-velocity-95th-centile-secondary-endpoint-duchenne-muscular-dystrophy_en.pdf
165. Viceconti M, Tome M, Dartee W, Knezevic I, Hernandez Penna S, Mazzà C, et al. On the use of wearable sensors as mobility biomarkers in the marketing authorization of new drugs: A regulatory perspective. *Frontiers in Medicine*. 2022;9. Available: <https://www.frontiersin.org/articles/10.3389/fmed.2022.996903>
166. European Medicine Agency. Letter of support for Mobilise-D digital mobility outcomes as monitoring biomarkers. 2020. Available: https://www.ema.europa.eu/en/documents/other/letter-support-mobilise-d-digital-mobility-outcomes-monitoring-biomarkers_en.pdf

167. European Medicine Agency. Letter of support for Mobilise-D digital mobility outcomes as monitoring biomarkers. 2021. Available: https://www.ema.europa.eu/en/documents/other/letter-support-mobilise-d-digital-mobility-outcomes-monitoring-biomarkers-follow_en.pdf
168. Food & Drug Administration. Requests for Feedback and Meetings for Medical Device Submissions: The Q-Submission Program. 2021.
169. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*. 2005;53: 695–699. doi:10.1111/j.1532-5415.2005.53221.x
170. Trojaniello D, Cereatti A, Della Croce U. Gait direction of progression changes using shank worn MIMUs. Validation on healthy and TBI subjects. 2014. doi:10.13140/2.1.1682.9443
171. Salis F, Bertuletti S, Caruso M, Croce UD, Mazzà C, Cereatti A. Multi-sensor integration and data fusion for enhancing gait assessment in and out of the laboratory. *Gait & Posture*. 2019;74: 34. doi:10.1016/j.gaitpost.2019.07.493
172. Bertuletti S, Della Croce U, Cereatti A. A wearable solution for accurate step detection based on the direct measurement of the inter-foot distance. *J Biomech*. 2019;84: 274–277. doi:10.1016/j.jbiomech.2018.12.039
173. Assessing the Credibility of Computational Modeling and Simulation in Medical Device Submissions. FDA; 2021 Dec.
174. Danneskiold-Samsøe B, Bartels EM, Bülow PM, Lund H, Stockmarr A, Holm CC, et al. Isokinetic and isometric muscle strength in a healthy population with special reference to age and gender. *Acta Physiologica*. 2009;197: 1–68. doi:10.1111/j.1748-1716.2009.02022.x
175. Clark BC, Manini TM. Sarcopenia \neq dynapenia. *J Gerontol A Biol Sci Med Sci*. 2008;63: 829–834. doi:10.1093/gerona/63.8.829
176. Shafiee G, Keshtkar A, Soltani A, Ahadi Z, Larijani B, Heshmat R. Prevalence of sarcopenia in the world: a systematic review and meta-analysis of general population studies. *J Diabetes Metab Disord*. 2017;16: 21. doi:10.1186/s40200-017-0302-x
177. Alfano LN, Lowes LP, Flanigan KM, Mendell JR. Correlation of knee strength to functional outcomes in becker muscular dystrophy. *Muscle & Nerve*. 2013;47: 550–554. doi:10.1002/mus.23660
178. Lo JH, U KP, Yiu T, Ong MT, Lee WY. Sarcopenia: Current treatments and new regenerative therapeutic approaches. *Journal of Orthopaedic Translation*. 2020;23: 38–52. doi:10.1016/j.jot.2020.04.002
179. Manini TM, Clark BC. Dynapenia and Aging: An Update. *The Journals of Gerontology: Series A*. 2012;67A: 28–40. doi:10.1093/gerona/glr010
180. Mitchell W, Atherton P, Williams J, Larvin M, Lund J, Narici M. Sarcopenia, Dynapenia, and the Impact of Advancing Age on Human Skeletal Muscle Size and Strength; a Quantitative Review. *Frontiers in Physiology*. 2012;3. Available: <https://www.frontiersin.org/articles/10.3389/fphys.2012.00260>
181. Cesari M, Pahor M, Lauretani F, Zamboni V, Bandinelli S, Bernabei R, et al. Skeletal Muscle and Mortality Results From the InCHIANTI Study. *The Journals of Gerontology: Series A*. 2009;64A: 377–384. doi:10.1093/gerona/gln031
182. Manini TM, Visser M, Won-Park S, Patel KV, Strotmeyer ES, Chen H, et al. Knee Extension Strength Cutpoints for Maintaining Mobility. *Journal of the American Geriatrics Society*. 2007;55:

451–457. doi:10.1111/j.1532-5415.2007.01087.x

183. Visser M, Simonsick EM, Colbert LH, Brach J, Rubin SM, Kritchevsky SB, et al. Type and Intensity of Activity and Risk of Mobility Limitation: The Mediating Role of Muscle Parameters. *Journal of the American Geriatrics Society*. 2005;53: 762–770. doi:10.1111/j.1532-5415.2005.53257.x
184. Beaudart C, McCloskey E, Bruyère O, Cesari M, Rolland Y, Rizzoli R, et al. Sarcopenia in daily practice: assessment and management. *BMC Geriatr*. 2016;16: 170. doi:10.1186/s12877-016-0349-4
185. Mau-Moeller A, Behrens M, Lindner T, Bader R, Bruhn S. Age-related changes in neuromuscular function of the quadriceps muscle in physically active adults. *Journal of Electromyography and Kinesiology*. 2013;23: 640–648. doi:10.1016/j.jelekin.2013.01.009
186. Jenkins NDM, Housh TJ, Palmer TB, Cochrane KC, Bergstrom HC, Johnson GO, et al. Relative differences in strength and power from slow to fast isokinetic velocities may reflect dynapenia. *Muscle & Nerve*. 2015;52: 120–130. doi:10.1002/mus.24505
187. Watanabe K, Kouzaki M, Ogawa M, Akima H, Moritani T. Relationships between muscle strength and multi-channel surface EMG parameters in eighty-eight elderly. *European Review of Aging and Physical Activity*. 2018;15: 3. doi:10.1186/s11556-018-0192-z
188. Clark BC, Taylor JL. Age-related changes in motor cortical properties and voluntary activation of skeletal muscle. *Curr Aging Sci*. 2011;4: 192–199. doi:10.2174/1874609811104030192
189. Pasco JA, Stuart AL, Holloway-Kew KL, Tembo MC, Sui SX, Anderson KB, et al. Lower-limb muscle strength: normative data from an observational population-based study. *BMC Musculoskeletal Disorders*. 2020;21. doi:10.1186/s12891-020-3098-7
190. Dulac M, Boutros GEH, Pion C, Barbat-Artigas S, Gouspillou G, Aubertin-Leheudre M. Is handgrip strength normalized to body weight a useful tool to identify dynapenia and functional incapacity in post-menopausal women? *Braz J Phys Ther*. 2016;20: 510–516. doi:10.1590/bjpt-rbf.2014.0184
191. Soares AV, Marcelino E, Maia KC, Borges NG. Relation between functional mobility and dynapenia in institutionalized frail elderly. *Einstein (Sao Paulo)*. 2017;15: 278–282. doi:10.1590/S1679-45082017AO3932
192. Santos EPRD, Silva CFR, Ohara DG, Matos AP, Pinto ACPN, Pegorari MS. Short Physical Performance Battery (SPPB) score as a discriminator of dynapenic abdominal obesity among community-dwelling older adults. *Geriatr Nurs*. 2021;42: 467–472. doi:10.1016/j.gerinurse.2021.02.016
193. Ribeiro JC, Duarte JG, Gomes GAO, Costa-Guarisco LP, de Jesus ITM, Nascimento CMC, et al. Associations between inflammatory markers and muscle strength in older adults according to the presence or absence of obesity. *Experimental Gerontology*. 2021;151: 111409. doi:10.1016/j.exger.2021.111409
194. Minetto MA, Ghigo E. Sarcopenia in Endocrine Disorders – The Iceberg or Its Tip? *Eur Endocrinol*. 2015;11: 41–42. doi:10.17925/EE.2015.11.01.41
195. Thanaj M, Bastly N, Whitcher B, Sorokin EP, Liu Y, Srinivasan R, et al. Precision MRI phenotyping of muscle volume and quality at a population scale. *Front Physiol*. 2024;15: 1288657. doi:10.3389/fphys.2024.1288657
196. Sinelnikov A, Qu C, Fetzer DT, Pelletier J-S, Dunn MA, Tsung A, et al. Measurement of skeletal muscle area: Comparison of CT and MR imaging. *European Journal of Radiology*. 2016;85: 1716–1721. doi:10.1016/j.ejrad.2016.07.006

197. Hong A, Liu JN, Gowd AK, Dhawan A, Amin NH. Reliability and Accuracy of MRI in Orthopedics: A Survey of Its Use and Perceived Limitations. *Clin Med Insights Arthritis Musculoskelet Disord.* 2019;12: 1179544119872972. doi:10.1177/1179544119872972
198. Hargreaves BA, Worters PW, Pauly KB, Pauly JM, Koch KM, Gold GE. Metal-Induced Artifacts in MRI. *American Journal of Roentgenology.* 2011;197: 547–555. doi:10.2214/AJR.11.7364
199. Munn Z, Moola S, Lisy K, Riitano D, Murphy F. Claustrophobia in magnetic resonance imaging: A systematic review and meta-analysis. *Radiography.* 2015;21: e59–e63. doi:10.1016/j.radi.2014.12.004
200. Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: A complex problem with many partial solutions. *Journal of Magnetic Resonance Imaging.* 2015;42: 887–901. doi:10.1002/jmri.24850
201. Owen G, Cronin J, Gill N, McNair P. Knee extensor stiffness and functional performance. *Physical Therapy in Sport.* 2005;6: 38–44. doi:10.1016/j.ptsp.2004.11.003
202. Fukunaga T, Johnson CD, Nicholas SJ, McHugh MP. Muscle hypotrophy, not inhibition, is responsible for quadriceps weakness during rehabilitation after anterior cruciate ligament reconstruction. *Knee Surg Sports Traumatol Arthrosc.* 2019;27: 573–579. doi:10.1007/s00167-018-5166-1
203. Şimşek D, Kirkaya İ, Onarici Güngör E, Soylu AR. Relationships Among Vertical Jumping Performance, EMG Activation, and Knee Extensor and Flexor Muscle Strength in Turkish Elite Male Volleyball Players. *Turkiye Klinikleri J Sports Sci.* 2016;8: 46–56. doi:10.5336/sportsci.2016-51433
204. Ettinger L, Weiss J, Shapiro M, Karduna A. Normalization to Maximal Voluntary Contraction is Influenced by Subacromial Pain. *Journal of Applied Biomechanics.* 2016;32: 433–440. doi:10.1123/jab.2015-0185
205. Ball N, Scurr J. An assessment of the reliability and standardisation of tests used to elicit reference muscular actions for electromyographical normalisation. *Journal of Electromyography and Kinesiology.* 2010;20: 81–88. doi:10.1016/j.jelekin.2008.09.004
206. Peacock B, Westers T, Walsh S, Nicholson K. Feedback and maximum voluntary contraction. *Ergonomics.* 1981;24: 223–228. doi:10.1080/00140138108559236
207. Tsiros MD, Grimshaw PN, Schield AJ, Buckley JD. Test-retest reliability of the Biodex System 4 Isokinetic Dynamometer for knee strength assessment in paediatric populations. *J Allied Health.* 2011;40: 115–119.
208. Pavone E, Moffat M. Isometric torque of the quadriceps femoris after concentric, eccentric and isometric training. *Arch Phys Med Rehabil.* 1985;66: 168–170.
209. Drouin JM, Valovich-mcLeod TC, Shultz SJ, Gansneder BM, Perrin DH. Reliability and validity of the Biodex system 3 pro isokinetic dynamometer velocity, torque and position measurements. *Eur J Appl Physiol.* 2004;91: 22–29. doi:10.1007/s00421-003-0933-0
210. Maffiuletti NA, Bizzini M, Desbrosses K, Babault N, Munzinger U. Reliability of knee extension and flexion measurements using the Con-Trex isokinetic dynamometer. *Clinical Physiology and Functional Imaging.* 2007;27: 346–353. doi:10.1111/j.1475-097X.2007.00758.x
211. Lund H, Søndergaard K, Zachariassen T, Christensen R, Bülow P, Henriksen M, et al. Learning effect of isokinetic measurements in healthy subjects, and reliability and comparability of Biodex and Lido dynamometers. *Clin Physiol Funct Imaging.* 2005;25: 75–82. doi:10.1111/j.1475-097X.2004.00593.x
212. Arampatzis A, Karamanidis K, De Monte G, Stafilidis S, Morey-Klapsing G, Brüggemann G-P.

- Differences between measured and resultant joint moments during voluntary and artificially elicited isometric knee extension contractions. *Clinical Biomechanics*. 2004;19: 277–283. doi:10.1016/j.clinbiomech.2003.11.011
213. Tsaopoulos DE, Baltzopoulos V, Richards PJ, Maganaris CN. Mechanical correction of dynamometer moment for the effects of segment motion during isometric knee-extension tests. *J Appl Physiol* (1985). 2011;111: 68–74. doi:10.1152/jappphysiol.00821.2010
 214. Stark T, Walker B, Phillips JK, Fejer R, Beck R. Hand-held Dynamometry Correlation With the Gold Standard Isokinetic Dynamometry: A Systematic Review. *PM&R*. 2011;3: 472–479. doi:10.1016/j.pmrj.2010.10.025
 215. Sarabon N, Kozinc Z, Bishop C, Maffiuletti NA. Factors influencing bilateral deficit and inter-limb asymmetry of maximal and explosive strength: motor task, outcome measure and muscle group. *Eur J Appl Physiol*. 2020;120: 1681–1688. doi:10.1007/s00421-020-04399-1
 216. Gagnon D, Nadeau S, Gravel D, Robert J, Bélanger D, Hilsenrath M. Reliability and Validity of Static Knee Strength Measurements Obtained With a Chair-Fixed Dynamometer in Subjects With Hip or Knee Arthroplasty. *Archives of Physical Medicine and Rehabilitation*. 2005;86: 1998–2008. doi:10.1016/j.apmr.2005.04.013
 217. Rodríguez-García WD, García-Castañeda L, Orea-Tejeda A, Mendoza-Núñez V, González-Islas DG, Santillán-Díaz C, et al. Handgrip strength: Reference values and its relationship with bioimpedance and anthropometric variables. *Clinical Nutrition ESPEN*. 2017;19: 54–58. doi:10.1016/j.clnesp.2017.01.010
 218. Lee S-C, Wu L-C, Chiang S-L, Lu L-H, Chen C-Y, Lin C-H, et al. Validating the Capability for Measuring Age-Related Changes in Grip-Force Strength Using a Digital Hand-Held Dynamometer in Healthy Young and Elderly Adults. *BioMed Research International*. 2020;2020: e6936879. doi:10.1155/2020/6936879
 219. Hogrel J-Y. Grip strength measured by high precision dynamometry in healthy subjects from 5 to 80 years. *BMC Musculoskelet Disord*. 2015;16: 139. doi:10.1186/s12891-015-0612-4
 220. Lupton-Smith A, Fourie K, Mazinyo A, Mokone M, Nxaba S, Morrow B. Measurement of hand grip strength: A cross-sectional study of two dynamometry devices. *S Afr J Physiother*. 2022;78: 1768. doi:10.4102/sajp.v78i1.1768
 221. De Luca CJ. The Use of Surface Electromyography in Biomechanics. *Journal of Applied Biomechanics*. 1997;13: 135–163. doi:10.1123/jab.13.2.135
 222. Hagberg M. The amplitude distribution of surface EMG in static and intermittent static muscular performance. *Europ J Appl Physiol*. 1979;40: 265–272. doi:10.1007/BF00421518
 223. Bray SR, Martin Ginis KA, Hicks AL, Woodgate J. Effects of self-regulatory strength depletion on muscular performance and EMG activation. *Psychophysiology*. 2008;45: 337–343. doi:10.1111/j.1469-8986.2007.00625.x
 224. Sturma A, Hrubby LA, Prahm C, Mayer JA, Aszmann OC. Rehabilitation of Upper Extremity Nerve Injuries Using Surface EMG Biofeedback: Protocols for Clinical Application. *Frontiers in Neuroscience*. 2018;12. Available: <https://www.frontiersin.org/articles/10.3389/fnins.2018.00906>
 225. Stålberg E, Dioszeghy P. Scanning EMG in normal muscle and in neuromuscular disorders. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*. 1991;81: 403–416. doi:10.1016/0168-5597(91)90048-3
 226. Subasi A. Classification of EMG signals using PSO optimized SVM for diagnosis of neuromuscular disorders. *Computers in Biology and Medicine*. 2013;43: 576–586.

doi:10.1016/j.compbiomed.2013.01.020

227. Rayegani SM, Raeissadat SA, Sedighipour L, Rezazadeh IM, Bahrami MH, Eliaspour D, et al. Effect of neurofeedback and electromyographic-biofeedback therapy on improving hand function in stroke patients. *Top Stroke Rehabil.* 2014;21: 137–151. doi:10.1310/tsr2102-137
228. Konrad P. *The ABC of EMG - A practical introduction to kinesiological electromyography. A practical introduction to kinesiological electromyography.* 2005;1.
229. Farina D, Merletti R, Enoka RM. The extraction of neural strategies from the surface EMG. *J Appl Physiol (1985).* 2004;96: 1486–1495. doi:10.1152/jappphysiol.01070.2003
230. Kamavuako EN, Rosenvang JC, Horup R, Jensen W, Farina D, Englehart KB. Surface Versus Untargeted Intramuscular EMG Based Classification of Simultaneous and Dynamically Changing Movements. *IEEE Transactions on Neural Systems and Rehabilitation Engineering.* 2013;21: 992–998. doi:10.1109/TNSRE.2013.2248750
231. Raez MBI, Hussain MS, Mohd-Yasin F. Techniques of EMG signal analysis: detection, processing, classification and applications. *Biol Proced Online.* 2006;8: 11–35. doi:10.1251/bpo115
232. Chowdhury RH, Reaz MBI, Ali MABM, Bakar AAA, Chellappan K, Chang TG. Surface Electromyography Signal Processing and Classification Techniques. *Sensors.* 2013;13: 12431–12466. doi:10.3390/s130912431
233. Merletti R, Cerone GL. Tutorial. Surface EMG detection, conditioning and pre-processing: Best practices. *Journal of Electromyography and Kinesiology.* 2020;54: 102440. doi:10.1016/j.jelekin.2020.102440
234. Hermens HJ. The state of the art on sensors and sensor placement procedures for surface electromyography : a proposal for sensor placement procedures ; deliverable of the SENIAM project. Roessingh Research and Development; 1997.
235. Cesarelli M, Bifulco P, Bracale M. Quadriceps muscles activation in anterior knee pain during isokinetic exercise. *Med Eng Phys.* 1999;21: 469–478. doi:10.1016/s1350-4533(99)00068-5
236. Manal K, Buchanan TS. An Electromyogram-Driven Musculoskeletal Model of the Knee to Predict in Vivo Joint Contact Forces During Normal and Novel Gait Patterns. *Journal of Biomechanical Engineering.* 2013;135. doi:10.1115/1.4023457
237. Sheffler LR, Chae J. Neuromuscular electrical stimulation in neurorehabilitation. *Muscle Nerve.* 2007;35: 562–590. doi:10.1002/mus.20758
238. Blazeovich AJ, Collins DF, Millet GY, Vaz MA, Maffiuletti NA. Enhancing Adaptations to Neuromuscular Electrical Stimulation Training Interventions. *Exerc Sport Sci Rev.* 2021;49: 244–252. doi:10.1249/JES.0000000000000264
239. Labanca L, Rocchi JE, Laudani L, Guitaldi R, Virgulti A, Mariani PP, et al. Neuromuscular Electrical Stimulation Superimposed on Movement Early after ACL Surgery. *Medicine & Science in Sports & Exercise.* 2018;50: 407. doi:10.1249/MSS.0000000000001462
240. Maffiuletti NA, Gondin J, Place N, Stevens-Lapsley J, Vivodtzev I, Minetto MA. Clinical Use of Neuromuscular Electrical Stimulation for Neuromuscular Rehabilitation: What Are We Overlooking? *Arch Phys Med Rehabil.* 2018;99: 806–812. doi:10.1016/j.apmr.2017.10.028
241. Vanderthommen M, Duchateau J. Electrical Stimulation as a Modality to Improve Performance of the Neuromuscular System. *Exercise and Sport Sciences Reviews.* 2007;35: 180. doi:10.1097/jes.0b013e318156e785
242. Borzuola R, Laudani L, Labanca L, Macaluso A. Superimposing neuromuscular electrical

- stimulation onto voluntary contractions to improve muscle strength and mass: A systematic review. *Eur J Sport Sci.* 2023;23: 1547–1559. doi:10.1080/17461391.2022.2104656
243. Osborne JO, Tallent J, Girard O, Marshall PW, Kidgell D, Buhmann R. Neuromuscular electrical stimulation during maximal voluntary contraction: a Delphi survey with expert consensus. *Eur J Appl Physiol.* 2023;123: 2203–2212. doi:10.1007/s00421-023-05232-1
 244. Glaviano NR, Saliba S. Can the Use of Neuromuscular Electrical Stimulation Be Improved to Optimize Quadriceps Strengthening? *Sports Health.* 2016;8: 79–85. doi:10.1177/1941738115618174
 245. Kyle UG, Bosaeus I, De Lorenzo AD, Deurenberg P, Elia M, Gómez JM, et al. Bioelectrical impedance analysis—part I: review of principles and methods. *Clinical Nutrition.* 2004;23: 1226–1243. doi:10.1016/j.clnu.2004.06.004
 246. Kushner RF. Bioelectrical impedance analysis: a review of principles and applications. *J Am Coll Nutr.* 1992;11: 199–209.
 247. Kyle UG, Bosaeus I, De Lorenzo AD, Deurenberg P, Elia M, Manuel Gómez J, et al. Bioelectrical impedance analysis—part II: utilization in clinical practice. *Clinical Nutrition.* 2004;23: 1430–1453. doi:10.1016/j.clnu.2004.09.012
 248. Jung H, Tanaka S, Tanaka R. Body Composition Characteristics of Community-Dwelling Older Adults With Dynapenia. *Front Nutr.* 2022;9: 827114. doi:10.3389/fnut.2022.827114
 249. Jewett JW, Serway RA. *Physics for Scientists and Engineers with Modern Physics.* 7th ed. Pacific Grove, California: Thomson Brooks/Cole; 2007.
 250. Guilak F. The slippery slope of arthritis. *Arthritis Rheum.* 2005;52: 1632–1633. doi:10.1002/art.21051
 251. Wu G, Siegler S, Allard P, Kirtley C, Leardini A, Rosenbaum D, et al. ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—part I: ankle, hip, and spine. *Journal of Biomechanics.* 2002;35: 543–548. doi:10.1016/S0021-9290(01)00222-6
 252. Lenhart RL, Kaiser J, Smith CR, Thelen DG. Prediction and Validation of Load-Dependent Behavior of the Tibiofemoral and Patellofemoral Joints During Movement. *Ann Biomed Eng.* 2015;43: 2675–2685. doi:10.1007/s10439-015-1326-3
 253. Huxley H, Hanson J. Changes in the cross-striations of muscle during contraction and stretch and their structural interpretation. *Nature.* 1954;173: 973–976. doi:10.1038/173973a0
 254. Röhrle O, Sprenger M, Schmitt S. A two-muscle, continuum-mechanical forward simulation of the upper limb. *Biomech Model Mechanobiol.* 2017;16: 743–762. doi:10.1007/s10237-016-0850-x
 255. Frey Law LA, Shields RK. Predicting human chronically paralyzed muscle force: a comparison of three mathematical models. *J Appl Physiol (1985).* 2006;100: 1027–1036. doi:10.1152/jappphysiol.00935.2005
 256. Caillet AH, Phillips A, Carty C, Farina D, Modenese L. Hill-type computational models of muscle-tendon actuators: a systematic review. 2022. doi:10.1101/2022.10.14.512218
 257. Thelen DG, Anderson FC, Delp SL. Generating dynamic simulations of movement using computed muscle control. *Journal of Biomechanics.* 2003;36: 321–328. doi:10.1016/S0021-9290(02)00432-3
 258. Millard M, Uchida T, Seth A, Delp SL. Flexing Computational Muscle: Modeling and Simulation of Musculotendon Dynamics. *J Biomech Eng.* 2013;135: 0210051–02100511. doi:10.1115/1.4023390

259. Seth A, Hicks JL, Uchida TK, Habib A, Dembia CL, Dunne JJ, et al. OpenSim: Simulating musculoskeletal dynamics and neuromuscular control to study human and animal movement. *PLOS Computational Biology*. 2018;14: e1006223. doi:10.1371/journal.pcbi.1006223
260. Rasmussen J, Vondrak V, Damsgaard M, de Zee M, Tørholm S, Dostal Z. The anybody project-computer analysis of the human body. *Biomechanics of Man*. 2002.
261. Scheys L, Spaepen A, Suetens P, Jonkers I. Calculated moment-arm and muscle-tendon lengths during gait differ substantially using MR based versus rescaled generic lower-limb musculoskeletal models. *Gait Posture*. 2008;28: 640–648. doi:10.1016/j.gaitpost.2008.04.010
262. Wagner DW, Stepanyan V, Shippen JM, Demers MS, Gibbons RS, Andrews BJ, et al. Consistency among musculoskeletal models: caveat utilitor. *Ann Biomed Eng*. 2013;41: 1787–1799. doi:10.1007/s10439-013-0843-1
263. Handsfield GG, Meyer CH, Abel MF, Blemker SS. Heterogeneity of muscle sizes in the lower limbs of children with cerebral palsy. *Muscle Nerve*. 2016;53: 933–945. doi:10.1002/mus.24972
264. Steele KM, DeMers MS, Schwartz MS, Delp SL. Compressive Tibiofemoral Force during Crouch Gait. *Gait Posture*. 2012;35: 556–560. doi:10.1016/j.gaitpost.2011.11.023
265. Handsfield GG, Meyer CH, Hart JM, Abel MF, Blemker SS. Relationships of 35 lower limb muscles to height and body mass quantified using MRI. *Journal of Biomechanics*. 2014;47: 631–638. doi:10.1016/j.jbiomech.2013.12.002
266. Luis I, Afschrift M, De Groote F, Gutierrez-Farewik EM. Evaluation of musculoskeletal models, scaling methods, and performance criteria for estimating muscle excitations and fiber lengths across walking speeds. *Frontiers in Bioengineering and Biotechnology*. 2022;10. Available: <https://www.frontiersin.org/articles/10.3389/fbioe.2022.1002731>
267. van der Krogt MM, Bar-On L, Kindt T, Desloovere K, Harlaar J. Neuro-musculoskeletal simulation of instrumented contracture and spasticity assessment in children with cerebral palsy. *Journal of NeuroEngineering and Rehabilitation*. 2016;13: 64. doi:10.1186/s12984-016-0170-5
268. van Veen B, Montefiori E, Modenese L, Mazzà C, Viceconti M. Muscle recruitment strategies can reduce joint loading during level walking. *Journal of Biomechanics*. 2019;97: 109368. doi:10.1016/j.jbiomech.2019.109368
269. Seireg A, Arvikar RJ. A mathematical model for evaluation of forces in lower extremities of the musculo-skeletal system. *J Biomech*. 1973;6: 313–326. doi:10.1016/0021-9290(73)90053-5
270. Lund ME, de Zee M, Andersen MS, Rasmussen J. On validation of multibody musculoskeletal models. *Proc Inst Mech Eng H*. 2012;226: 82–94. doi:10.1177/0954411911431516
271. Hicks JL, Uchida TK, Seth A, Rajagopal A, Delp SL. Is My Model Good Enough? Best Practices for Verification and Validation of Musculoskeletal Models and Simulations of Movement. *J Biomech Eng*. 2015;137: 0209051–02090524. doi:10.1115/1.4029304
272. Henninger HB, Reese SP, Anderson AE, Weiss JA. Validation of computational models in biomechanics. *Proc Inst Mech Eng H*. 2010;224: 801–812. doi:10.1243/09544119JEIM649
273. Rooks NB, Schneider MTY, Erdemir A, Halloran JP, Laz PJ, Shelburne KB, et al. Deciphering the “Art” in Modeling and Simulation of the Knee Joint: Variations in Model Development. *J Biomech Eng*. 2021;143: 061002. doi:10.1115/1.4050028
274. Davico G, Bottin F, Di Martino A, Castafaro V, Baruffaldi F, Faldini C, et al. Intra-operator Repeatability of Manual Segmentations of the Hip Muscles on Clinical Magnetic Resonance Images. *J Digit Imaging*. 2023;36: 143–152. doi:10.1007/s10278-022-00700-0

275. Princelle D, Davico G, Viceconti M. Comparative validation of two patient-specific modelling pipelines for predicting knee joint forces during level walking. *J Biomech.* 2023;159: 111758. doi:10.1016/j.jbiomech.2023.111758
276. Meise HF, Lüpke M, Seifert H, Harms O. Development of a three-dimensional computer model of the canine pelvic limb including cruciate ligaments to simulate movement. *Research in Veterinary Science.* 2021;136: 430–443. doi:10.1016/j.rvsc.2021.03.015
277. Ascani D, Mazzà C, De Lollis A, Bernardoni M, Viceconti M. A procedure to estimate the origins and the insertions of the knee ligaments from computed tomography images. *Journal of Biomechanics.* 2015;48: 233–237. doi:10.1016/j.jbiomech.2014.11.041
278. Viceconti M, Ascani D, Mazzà C. Pre-operative prediction of soft tissue balancing in knee arthroplasty part 1: Effect of surgical parameters during level walking. *Journal of Orthopaedic Research.* 2019;37: 1537–1545. doi:10.1002/jor.24289
279. Lerner ZF, DeMers MS, Delp SL, Browning RC. How tibiofemoral alignment and contact locations affect predictions of medial and lateral tibiofemoral contact forces. *Journal of Biomechanics.* 2015;48: 644–650. doi:10.1016/j.jbiomech.2014.12.049
280. Modenese L, Ceseracciu E, Reggiani M, Lloyd DG. Estimation of musculotendon parameters for scaled and subject specific musculoskeletal models using an optimization technique. *Journal of Biomechanics.* 2016;49: 141–148. doi:10.1016/j.jbiomech.2015.11.006
281. Bennett KJ, Pizzolato C, Martelli S, Bahl JS, Sivakumar A, Atkins GJ, et al. EMG-Informed Neuromusculoskeletal Models Accurately Predict Knee Loading Measured Using Instrumented Implants. *IEEE Transactions on Biomedical Engineering.* 2022;69: 2268–2275. doi:10.1109/TBME.2022.3141067
282. Falisse A, Van Rossom S, Jonkers I, De Groot F. EMG-Driven Optimal Estimation of Subject-SPECIFIC Hill Model Muscle–Tendon Parameters of the Knee Joint Actuators. *IEEE Transactions on Biomedical Engineering.* 2017;64: 2253–2262. doi:10.1109/TBME.2016.2630009
283. Davico G, Lloyd DG, Carty CP, Killen BA, Devaprakash D, Pizzolato C. Multi-level personalization of neuromusculoskeletal models to estimate physiologically plausible knee joint contact forces in children. *Biomech Model Mechanobiol.* 2022;21: 1873–1886. doi:10.1007/s10237-022-01626-w
284. Hoang HX, Diamond LE, Lloyd DG, Pizzolato C. A calibrated EMG-informed neuromusculoskeletal model can appropriately account for muscle co-contraction in the estimation of hip joint contact forces in people with hip osteoarthritis. *Journal of Biomechanics.* 2019;83: 134–142. doi:10.1016/j.jbiomech.2018.11.042
285. Thompson BJ. Influence of signal filtering and sample rate on isometric torque – time parameters using a traditional isokinetic dynamometer. *Journal of Biomechanics.* 2019;83: 235–242. doi:10.1016/j.jbiomech.2018.12.006
286. Kooistra RD, de Ruitter CJ, de Haan A. Conventionally assessed voluntary activation does not represent relative voluntary torque production. *Eur J Appl Physiol.* 2007;100: 309–320. doi:10.1007/s00421-007-0425-8
287. Babault N, Pousson M, Michaut A, Ballay Y, Hoecke JV. EMG activity and voluntary activation during knee-extensor concentric torque generation. *Eur J Appl Physiol.* 2002;86: 541–547. doi:10.1007/s00421-002-0579-3
288. Cignoni P, Callieri M, Corsini M, Dellepiane M, Ganovelli F, Ranzuglia G. MeshLab: an Open-Source Mesh Processing Tool. The Eurographics Association; 2008. doi:10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136

289. Dumas R, Aissaoui R, Mitton D, Skalli W, de Guise JA. Personalized body segment parameters from biplanar low-dose radiography. *IEEE Transactions on Biomedical Engineering*. 2005;52: 1756–1763. doi:10.1109/TBME.2005.855711
290. White DR, Woodard HQ, Hammond SM. Average soft-tissue and bone models for use in radiation dosimetry. *BJR*. 1987;60: 907–913. doi:10.1259/0007-1285-60-717-907
291. Zajac FE. Muscle and tendon: properties, models, scaling, and application to biomechanics and motor control. *Crit Rev Biomed Eng*. 1989;17: 359–411.
292. Lloyd DG, Buchanan TS. A Model of Load Sharing Between Muscles and Soft Tissues at the Human Knee During Static Tasks. *Journal of Biomechanical Engineering*. 1996;118: 367–376. doi:10.1115/1.2796019
293. Lloyd DG, Besier TF. An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo. *Journal of Biomechanics*. 2003;36: 765–776. doi:10.1016/S0021-9290(03)00010-1
294. Boccia G, Martinez-Valdes E, Negro F, Rainoldi A, Falla D. Motor unit discharge rate and the estimated synaptic input to the vasti muscles is higher in open compared with closed kinetic chain exercise. *Journal of Applied Physiology*. 2019;127: 950–958. doi:10.1152/jappphysiol.00310.2019
295. Martinez-Valdes E, Negro F, Falla D, De Nunzio AM, Farina D. Surface electromyographic amplitude does not identify differences in neural drive to synergistic muscles. *Journal of Applied Physiology*. 2018;124: 1071–1079. doi:10.1152/jappphysiol.01115.2017
296. Sartori M, Reggiani M, Farina D, Lloyd DG. EMG-Driven Forward-Dynamic Estimation of Muscle Force and Joint Moment about Multiple Degrees of Freedom in the Human Lower Extremity. *PLOS ONE*. 2012;7: e52618. doi:10.1371/journal.pone.0052618
297. Mellor R, Hodges P. Motor unit synchronization between medial and lateral vasti muscles. *Clinical Neurophysiology*. 2005;116: 1585–1595. doi:10.1016/j.clinph.2005.04.004
298. Petrone N, Tregnaghi D, Nardon M, Marcolin G. Musculoskeletal Simulation of Isokinetic Exercises: A Biomechanical and Electromyographical Pilot Study. *Procedia Engineering*. 2015;112: 250–255. doi:10.1016/j.proeng.2015.07.210
299. Crowninshield RD, Brand RA. A physiologically based criterion of muscle force prediction in locomotion. *Journal of Biomechanics*. 1981;14: 793–801. doi:10.1016/0021-9290(81)90035-X
300. Yoon TS, Park DS, Kang SW, Chun SI, Shin JS. Isometric and isokinetic torque curves at the knee joint. *Yonsei Medical Journal*. 1991;32: 33–43. doi:10.3349/ymj.1991.32.1.33
301. Bohannon RW, Kindig J, Sabo G, Duni AE, Cram P. Isometric knee extension force measured using a handheld dynamometer with and without belt-stabilization. *Physiotherapy Theory and Practice*. 2012;28: 562–568. doi:10.3109/09593985.2011.640385
302. Murray MP, Duthie EH, Gambert SR, Sepic SB, Mollinger LA. Age-Related Differences in Knee Muscle Strength in Normal Women. *J Gerontol*. 1985;40: 275–280. doi:10.1093/geronj/40.3.275
303. Redl C, Gfoehler M, Pandy MG. Sensitivity of muscle force estimates to variations in muscle–tendon properties. *Human Movement Science*. 2007;26: 306–319. doi:10.1016/j.humov.2007.01.008
304. Sartori M, Rubenson J, Lloyd DG, Farina D, Panizzolo FA. Subject-Specificity via 3D Ultrasound and Personalized Musculoskeletal Modeling. In: Ibáñez J, González-Vargas J, Azorín JM, Akay M, Pons JL, editors. *Converging Clinical and Engineering Research on Neurorehabilitation II*. Cham: Springer International Publishing; 2017. pp. 639–642. doi:10.1007/978-3-319-46669-9_105
305. Gogeoascoechea A, Ornelas-Kobayashi R, Yavuz US, Sartori M. Characterization of Motor Unit

- Firing and Twitch Properties for Decoding Musculoskeletal Force in the Human Ankle Joint In Vivo. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2023;31: 4040–4050. doi:10.1109/TNSRE.2023.3319959
306. Sartori M, Yavuz UŞ, Farina D. In Vivo Neuromechanics: Decoding Causal Motor Neuron Behavior with Resulting Musculoskeletal Function. *Sci Rep*. 2017;7: 13465. doi:10.1038/s41598-017-13766-6
307. Erdemir A, Besier TF, Halloran JP, Imhauser CW, Laz PJ, Morrison TM, et al. Deciphering the “Art” in Modeling and Simulation of the Knee Joint: Overall Strategy. *J Biomech Eng*. 2019;141: 0710021–07100210. doi:10.1115/1.4043346
308. Andreassen TE, Laz PJ, Erdemir A, Besier TF, Halloran JP, Imhauser CW, et al. Deciphering the “Art” in Modeling and Simulation of the Knee Joint: Assessing Model Calibration Workflows and Outcomes. *J Biomech Eng*. 2023;145: 121008. doi:10.1115/1.4063627
309. U.S. Food & Drug Administration. CDRH Regulatory Science Priorities (FY2017).
310. Viceconti M, Taddei F, Cristofolini L, Martelli S, Falcinelli C, Schileo E. Are spontaneous fractures possible? An example of clinical application for personalised, multiscale neuro-musculo-skeletal modelling. *Journal of Biomechanics*. 2012;45: 421–426. doi:10.1016/j.jbiomech.2011.11.048

Acknowledgements

As I conclude my doctoral journey, I am deeply grateful to those who have made this path possible.

First and foremost, I must express my profound gratitude to my supervisor, Prof. Viceconti, whose expert guidance, support, and trust have been the cornerstone of my academic journey. His wisdom and deep knowledge of the subject matter, along with his constant encouragement, have not only shaped my approach to research but have also greatly contributed to my personal growth and resilience in facing academic challenges. Furthermore, his supportive nature made him more than just a supervisor; he has been a trusted advisor who genuinely cared about my development as a scholar and as a person.

I also extend my heartfelt thanks to my co-supervisor, Dr. Giorgio Davico, for his invaluable support and constant readiness to assist. His thoughtful advice and attentive listening have significantly enriched my research experience. His ability to provide fresh perspectives has been instrumental in enhancing the quality of my work, adding a valuable dimension to my academic endeavors.

A special thanks goes to all the members of the Medical Technology Laboratory. The collaboration, support, and inspiration that I have received from each of you have greatly enhanced my doctoral experience. I am thankful for every engaging discussion, for the assistance in times of need, and, above all, for the friendships that have blossomed.

I want to extend my heartfelt thanks to all my friends for their incredible support, for distracting me, and for all the laughter we shared. Your presence has made this journey lighter and more enjoyable.

A huge thank you to my entire family, who have always stood by me, consistently supporting and encouraging me. Your belief in my choices and aspirations has been the basis of my strength and perseverance.

Finally, I owe a huge thanks to Rrok, who is once again by my side, holding my hand as I reach this new milestone. To you, my foremost supporter, in whose eyes I see the respect you have for me. To you, in whose arms I have found refuge when I needed it most. And to our little Sofia, who made me a mother and helped keep me awake at nights, making it possible for me to complete this thesis on time. Your love and presence have been my greatest motivators. Thank you for being my family, my inspiration, and my heart.

The contributions of all of you have not only made this journey possible but also deeply fulfilling. Thank you sincerely.