ESCOLA POLITÉCNICA

PROGRAMA DE PÓS-GRADUAÇÃO

MESTRADO EM ENGENHARIA ELÉTRICA

ANA MARIA BENDER SEIDENFUSS DAS NEVES

# STATISTICAL PARAMETRIC MAPPING OF GROUND REACTION FORCES DURING WALKING OF AGED ADULTS

Porto Alegre

2021

PÓS-GRADUAÇÃO - *STRICTO SENSU*

Pontifícia Universidade Católica
do Rio Grande do Sul

ANA MARIA BENDER SEIDENFUSS DAS NEVES

# STATISTICAL PARAMETRIC MAPPING OF GROUND REACTION FORCE DURING WALKING OF AGED ADULTS

Master's degree dissertation presented to the Graduate Program in Electrical Engineering of the Pontifical Catholic University of Rio Grande do Sul, as requisite to obtain the title of master's degree in Electrical Engineering.

Concentration Area: Signals, Systems and Information Technology. Research Line: Biomedical Engineering.

Advisor: PhD. Rafael Reimann Baptista
Co-Advisor: PhD Marcus Fraga Vieira

Porto Alegre
2021

# Ficha Catalográfica

D229s das Neves, Ana Maria Bender Seidenfuss

Statistical Parametric Mapping of Ground Reaction Forces during Walking of Aged Adults / Ana Maria Bender Seidenfuss das Neves. – 2021.
123.
Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia Elétrica, PUCRS.

Orientador: Prof. Dr. Rafael Reimann Baptista.
Co-orientador: Prof. Dr. Marcus Fraga Vieira.

1. Marcha. 2. Caminhada. 3. Idosos. 4. Mapeamento Estatístico Paramétrico (SPM). 5. Processamento de Dados. I. Baptista, Rafael Reimann. II. Vieira, Marcus Fraga. III. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS
com os dados fornecidos pelo(a) autor(a).
Bibliotecária responsável: Clarissa Jesinska Selbach CRB-10/2051

Pontifícia Universidade Católica do Rio Grande do Sul
ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA - PPGEE

# STATISTICAL PARAMETRIC MAPPING OF GROUND REACTION FORCE DURING WALKING OF AGED ADULTS

**CANDIDATA: ANA MARIA BENDER SEIDENFUSS DAS NEVES**

Esta Dissertação de Mestrado foi julgada para obtenção do título de MESTRE EM ENGENHARIA ELÉTRICA e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Pontifícia Universidade Católica do Rio Grande do Sul.

**DR. RAFAEL REIMANN BAPTISTA - ORIENTADOR**

**DR. MARCUS FRAGA VIEIRA - COORIENTADOR**

**BANCA EXAMINADORA**

**DRA. DENISE PASCHOAL SOARES - INSTITUTO PIAGET DE PORTUGAL**

**DRA. ANA MARIA MARQUES DA SILVA - PPGEE - PUCRS**

PUCRS
Av. Ipiranga, 6681 - Predio 32 - Sala 507 | CEP 90619-900 | Porto Alegre, RS - Brasil
Fone: (51) 3320-3540 | E-mail: engenharia.pg.eletrica@pucrs.br | www.pucrs.br/politecnica

## ACKNOWLEDGEMENTS

First, I would like to dedicate this work to my mother, Ana Paula Bender Seidenfuss, who have abdicated a great part of her life to raise me and my siblings, despite of any uncertainty of the future, she has always been there showing me what respect and love feels like. I would like to thank my father, Zilmo Alves das Neves for the support, understanding of my absence and for believing in me. I would like to also thank my siblings for the support and help with our parents while I had been dealing with the development of this work: Daiane, Valquiria, Julia, Vitor, Guilherme, Daniel, Jéssica Larissa, Paulo N., Marcos, Artur, and Paulo César despite of our differences you will always be important to me.

I would like to thank and dedicate this work to Rosângela Kuball da Silveira, she was the first person to show me what a monography was, I was fifteen years old and as she was struggling with her computer, she asked me to help digitating and formatting ABNT rules with her. Besides the digital struggle, she would write and talk about history and sociology effortlessly, and I wanted to be like that. She told me and I have never forgotten does not matter what you do, if someone asks you to do it, you go and do the best you can! I have been trying to keep that in mind. She told me I could study. Rosângela did a social work there, she told someone who had never thought of a different future that she could transform her own reality, she never said it would be an easy task, and I accepted the challenge, thank you for sharing your stories, history and example.

I would like to thank my friends, who have been listening and encouraging me through the achievement of my objectives. I would like to thank my dearest class and laboratory mates: Andréia, Morgana, Vandressa, Samantha and Thiago, being beside you have been a great experience of mutual respect, support and sharing of knowledge, technical and professional advice. Rodrigo for his important technical sharings during classes and group meetings, as well as both Leandro Giacomazzi and Leandro Sousa who have had patience to explain the main features of both mechanics and data processing that they had been developing and applying lead by Andréia Aires and Professor Rafael.

I would like to thank my professors and teachers who had accepted and sheltered me with my limitations, and oriented me toward my personal and professional development. I would like to thank all people who have worked towards affirmative

policies to make education more accessible. My advisor, professor Rafael Reimann Baptista who has been open to talk, discuss and conduct science in a multidisciplinary and ethical way. Minds who are open to solve problems and spread the word, so more people may have access to the knowledge produced in academy. Thank you for inviting me into this great adventure, thanks for your encouragement, and your ethical way of conducting your work. I would like to thank Professor Marcus Fraga Vieira for his advice and for sharing his knowledge on both engineering and physical education areas, which have greatly improved and made possible for me to accomplish this work. I would like to thank Professor Ana Maria Marques, for believing in me, and being a role model after, among and before many other researchers who are working towards inclusive and safe work environments.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nivel Superior – Brasil (CAPES) – Finance Code 001.

**RESUMO**

*O mapeamento estatístico paramétrico (SPM) atua ao longo da janela temporal de aquisição e detecta mudanças ao longo das dimensões de um campo contínuo n-dimensional. Objetiva-se investigar a usabilidade do SPM em dados experimentais de forças de reação do solo (FRS) durante a caminhada de idosas assim como contribuir para a expansão do uso do método na biomecânica.*

*Para tal realizou-se o estudo e execução de uma etapa preliminar de preparação de dados, seguido da análise tradicional com análise de pontos discretos e da análise contínua (SPM), e por fim a comparação dos resultados entre os métodos e entre os grupos etários (jovens, adultas e idosas). Dados experimentais de dois laboratórios foram incluídos neste estudo. Concluindo que é possível aplicar o método de análise continua em dados experimentais de FRS da marcha desde que se garanta uma boa preparação dos dados, bem como a expansão dos resultados numéricos para aplicações que auxiliem os profissionais das áreas de análise do movimento.*

*Os métodos apresentaram conclusões convergentes, entretanto o método contínuo expande a análise, resultando em uma gama maior de informação a partir de um teste único. Observou-se uma redução na amplitude da FRS para as idosas do conjunto de dados interno (Laboratório de Avaliação e Pesquisa em Atividade Física (LAPAFI - PUCRS)) quando comparado aos grupos de jovens e adultas do conjunto externo de dados (GaitRec) durante as fases de desaceleração e aceleração do apoio, e um aumento na fase de apoio simples (apoio médio), nos levando a acreditar que essas participantes confiam mais no grupo de músculos estabilizadores durante a marcha. Faz-se necessária a ampliação da aplicação do método e de suas dependências (aquisição, preparação, e gerenciamento de dados por exemplo).*

*Palavras-chave: marcha, caminhada, idosos, mapeamento estatístico paramétrico (SPM), processamento de dados.*

**ABSTRACT**

*Statistical parametric mapping (SPM) acts along the acquisition time window and detects changes along the dimensions of a continuous n-dimensional field. The objective is to investigate the usability of the SPM in experimental data of ground reaction forces (GRF) during walking in elderly women, as well as to contribute to the expansion of the use of the method in biomechanics.*

*For this, the study and execution of a preliminary stage of data preparation was carried out, followed by the traditional analysis with analysis of discrete points and continuous analysis (SPM), and finally the comparison of results between the methods and between the age groups (young, adult and elderly). Experimental data from two laboratories were included in this study. Concluding that it is possible to apply the method of continuous analysis in experimental data of GRF of the gait after a good preparation of the data is guaranteed, as well as the expansion of the numerical results for applications that help professionals in the areas of movement analysis.*

*The methods presented convergent conclusions; however the continuous method expands the analysis, resulting in a greater range of information from a single test. There was a reduction in the amplitude of the GRF for elderly women in the internal data set (Laboratory of Assessment and Research in Physical Activity (LAPAFI - PUCRS)) when compared to groups of young people and adults in the external data set (GaitRec) during the phases of deceleration and acceleration of stance, and an increase in the phase of simple stance (medium stance), leading us to believe that these participants rely more on the stabilizing muscle group during gait. It is necessary to expand the application of the method and its dependencies (acquisition, preparation, and data management, for example).*

*Keywords: gait, walk, elderly, parametric statistical mapping (SPM), data processing.*

## LIST OF FIGURES

## LIST OF TABLES

**CONTENTS**

# 1. INTRODUCTION

With the twentieth century's arrival, physicians started relying on other sources of information besides their education and daily instruments and we have seen an accelerated rise in the use of technology to aid medical decision both in diagnosis and treatment, and many hospitals became institutions of research and technolog. Biomedical engineering became a recognized profession as the effect of efforts of professionals from chemistry, physics, mechanical engineering, and electrical engineering working in conjunction with the medical field. As a result, medical technology advanced more in the twentieth century than it had in the rest of history (ENDERLE, 2012).

Among the variety of instrumentation systems available to quantify medical variables there are those dedicated to the analysis of locomotion. Locomotion includes a variety of types and is defined as the ability to move from one place to other. We are going to focus on bipedal legged terrestrial locomotion, more specifically on walking gait. The analysis of locomotion involves efforts on both internal and external aspects of the movement. Internal aspects are related to the physical experience of different locations of the body, such as internal forces on joints, electrical signals activities sent from the motor cortex to muscles, and acceleration of body segments. External aspects are related to the environment in which the movement occurs for example the type of shoes, the use of prothesis and the type of surface on which the movement occurs. In terms of energy expenditure by muscles, internal work is done to move the limbs relative to the center of mass while external work is done to move the center of mass of the body relative to the surrounding (CAVAGNA, 2017).

The gait cycle is a periodic phenomenon. It begins at a one-foot strike and ends at the consecutive foot strike at the same side, this is also called stride (BAKER, 2013). Measurements of motion and the forces underlying that motion are fundamental to biomechanical experimentation (PATAKY; ROBINSON; VANRENTERGHEM, 2013). Gait kinetics is the measure of the forces and moments (torques) that cause or restrict movement. There are extrinsic (external), and intrinsic (internal) joint torques (BEGG; WYTCH; MAJOR, 1989; SLOOT; VAN DER KROGT, 2018). The most common method used is to measure the reaction force of the foot floor using a force platform. These force platforms can measure forces in three orthogonal directions (one vertical and two horizontal) (BEGG; WYTCH; MAJOR, 1989). Ground reaction forces (GRF)

influence balance in the lateral median plane (ML), while the effects on accelerating and decelerating movement in the anterior-posterior (AP) plane. The vertical component (V) GRF acts perpendicular to the ground against the gravitational force (F. VAVERKA, 2015).

Usually, GRF and kinematic trajectories are sampled at frequencies that can result in hundreds or thousands of values per recording. These are often plotted as one-dimensional $1D$ biomechanical continuous or step-wise continuous $1D$ trajectories curves (PATAKY, 2012). Or $1D$ scalar trajectories $y_i(q)$, where $i$ represents a particular physical body, joint, axis or direction, and $q$ represents $1D$ time or space (PATAKY; ROBINSON; VANRENTERGHEM, 2013). Since these trajectories can be complex, it can be difficult to objectively specify an a priori method for analyzing them (PATAKY; VANRENTERGHEM; ROBINSON, 2015).

These curves are usually temporal or spatial normalized to a scale from 0% to 100%. Experimental research performs several measurements of $y_i(q)$ followed by registration, then mean and variance curves are computed, but traditionally hypothesis testing is conducted in a relatively small set of discrete values, such as values at local optima (maxima and minima) (PATAKY, 2012). This reduction to a few discrete values is computationally convenient but is not strictly necessary. The statistical parametric mapping (SPM) is an alternative to conduct statistical hypothesis testing continuously and to mitigate bias sources (PATAKY, 2012; PATAKY; ROBINSON; VANRENTERGHEM, 2013).

Many studies use an ad-hoc approach, which consists of visualizing the trajectories and then extracting some summary scalar, which was not specified before the experiment to be tested statistically (PATAKY; VANRENTERGHEM; ROBINSON, 2015). It is a non-trivial challenge to employ statistical methods that are consistent with one's null hypothesis. Directed null hypotheses and non-directed null-hypothesis have some limitations such as post-hoc regional focus bias and inter-component covariance bias (PATAKY; ROBINSON; VANRENTERGHEM, 2013). The bias is explained as all statistical analyses require a model of randomness. It is generally biased to test a (trajectory) $1D$ hypothesis using a (point) $0D$ model. A randomness model may separately be categorized as either parametric or non-parametric (PATAKY; VANRENTERGHEM; ROBINSON, 2015). SPM is an n-dimensional $nD$ methodology to analyze smooth continuum changes associated with experimental intervention. SPM has two main advantages over the extracting discrete optima values approach, which

are (i) statistical results are presented directly in the original sampling space, with spatiotemporal biomechanical content immediately apparent and (ii) no need for assumptions regarding the spatiotemporal foci of signals (PATAKY, 2012).

## 1.1. **Research Problem**

Can we guarantee experimental GRF data quality using pre-processing techniques? Can we apply SPM on biomechanics experimental studies, what are the limitations and challenges involved, and does SPM and traditional approach conclusions diverge? Are there differences on GRF experimental data between age groups of healthy women?

## 1.2. **General Objective**

The main goal of this research is to investigate the usability and methodological limitations and challenges of SPM application on experimental GRF data of human walking.

### 1.2.1. Specific Objectives

a) To develop a pre-processing routine to assure data quality and reproducibility.
b) To compare traditional (discrete) and SPM (continuous) methods applied to GRF data.
c) To compare GRF of different age groups of healthy women.

## 1.3. **Hypothesis**

a. It is possible to improve quality of experimental GRF data using pre-processing techniques.
b. It is possible to apply SPM and traditional analysis on experimental GRF data and that their results may lead to different conclusions.
c. There are differences on GRF stance phase between age groups: young, adult, and older adult which can be associated to the gait events and their physiological aspects.

# 2. STATE OF THE ART: GROUND REACTION FORCE (GRF) AND STATISTICAL PARAMETRIC MAPPING (SPM)

## 2.1. Methods

The literature research for SPM of the GRF of human gait analysis was conducted using the following keywords on three databases: (i) Scopus: 33 articles; (ii) Pubmed 13 articles; (iii) Web of Science: 31 articles. The terms used were "ground reaction force" AND "statistical parametric mapping". Only articles were included. These results were extracted into '.text' files and a total of 77 articles were inserted into an excel datasheet (Microsoft Office 360) and organized according to the following information: author, title, year, DOI, abstract, document type and source of origin. Forty-two duplicated studies were excluded through visual inspection of title matching, after that 35 titles and abstracts were read to minimize undesired exclusion of relevant work. Twenty-five studies did not meet our inclusion criteria, resulting in 10 articles for further investigation, from those 10 studies were included in the systematic review for critical analysis.

No limitation of the period was established, and the source research resulted in the following periods: Scopus (2012-2019), PubMed (2015-2019), and Web of Science: (2012-2019). The inclusion criteria were walking studies measuring GRF parameters and applying SPM methods. The research exclusion criteria were studies exclusively related to other physical activities rather than walking, studies with only kinematic parameters, jumping tasks, studies that specifically involved subjects with injuries or gait disorders, and animal studies. Only studies in English were included. There were no inclusion/exclusion criteria for publication stage, funding sponsors, country/territory, and affiliation. All authors were considered for abstract and title inspection. *Figure 1* presents the outcomes of research and filtering from the aforementioned databases.

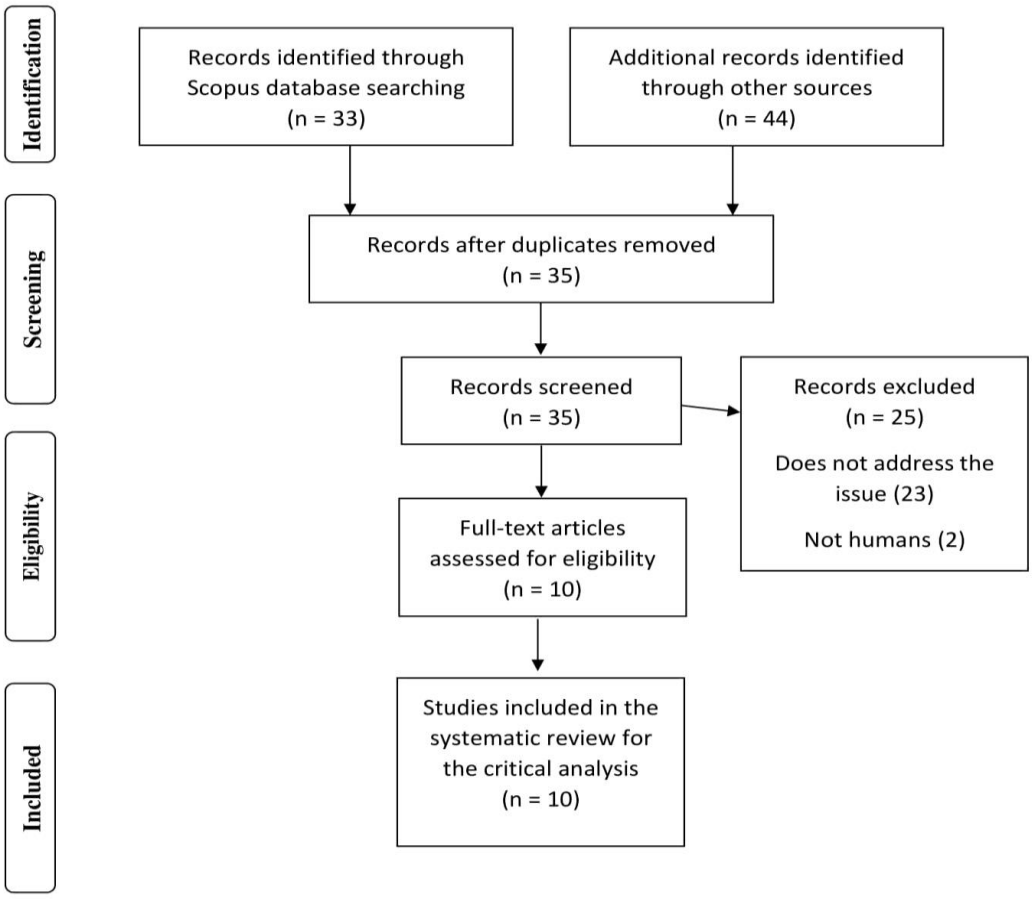



*Figure 1. Flow diagram of study selection. Source: The author, 2019.*

## 2.2. **Summary**

In 2012, Pataky published a study introducing 'SPM1D', a free and open-source software package for $1D$ SPM with basic pre-processing functionality and SPM computations in Python. The study also describes SPM procedures and SPM1D implementation in details followed by three SPM1D demonstrations using experimental datasets (Pataky, 2012). In 2013, SPM and Random Field Theory (RFT) were presented to conduct formalized hypothesis testing on publicly available biomechanical vector field datasets. Using mathematical arguments and logical interpretations of the original data, the authors concluded that scalar extraction constitutes a biased approach to non-directed hypothesis testing and that SPM overcomes these biases (PATAKY; ROBINSON; VANRENTERGHEM, 2013).

A systematic comparison of $0D$ vs. $1D$ procedures, parametric vs. non-parametric results and confidence interval (CI) vs. hypothesis testing elucidates the theoretical framework of $0D$ vs. $1D$ statistical procedures. Specifically, the choice of the procedure

($0D$ vs. $1D$ ) is statistically more relevant than choosing parametric vs. non-parametric procedures because greater differences in $0D$ vs. $1D$ results were observed when compared to parametric vs. non-parametric results; In contrast to $1D$ CIs which are complex and non-generalizable, $1D$ hypothesis testing results can be conducted consistently across all experimental designs (PATAKY; VANRENTERGHEM; ROBINSON, 2015).

In 2015, the influence of footwear and soft surface on gait initiation (GI) was investigated using the center of pressure (COP) and GRF time series during each gait initiation phase in three conditions (barefoot, barefoot with foam surface and wearing shoes). COP path, mean COP velocity, and force peaks were collected to identify differences between conditions and literature. Force impulse x COP integral to analyze COP displacement in body weight transfer between swing foot and support foot. To capture features of the entire COP time series instead of a few discrete variables, they conducted a vector analysis using SPM methods. The authors concluded that the use of footwear changes COP and GRFs behavior, and SPM analysis of COP time series also reveals significant differences in wearing shoes (VIEIRA; SACCO; NORA; ROSENBAUM *et al.*, 2015).

Also, in 2015 two statistical methods were used to investigate whether gender association is justified when assessing GRF while walking. The authors used two techniques to test the null hypothesis of equal mean three-component GRF waveforms between sexes: (i) SPM to analyze the entire three-component waveform with a single two-sample Hotelling $T^2$ test. The study suggests the advantages of using SPM to justify grouping because it conducts only a single waveform level test and therefore yields an unambiguous conclusion regarding data grouping. (ii) Approach by analyzing the first and second vertical GRF peaks with two separate t-tests for two samples, which produced comparatively ambiguous results, because it is possible to choose arbitrary metrics that yield multiple and potentially conflicting results. The authors suggest that researchers and clinicians consider the entire gait waveform and gender specificity when analyzing GRF data (CASTRO; PATAKY; SOLE; VILAS-BOAS, 2015).

In 2016, one study explored public datasets from a variety of experimental tasks and data modalities. Based on the temporal smoothness of the data, the authors quantified the false-positive rates expected in real $1D$ biomechanical datasets when employing $0D$ statistical inference, using a simplified experimental design: a two-

sample t-test. The key theoretical concept they shared was that two parameters, mean (μ) and standard deviation (σ), describe the $0D$ Gaussian behavior, and that only adding one parameter, $1D$ smoothness (FWHM), is necessary to describe the behavior of Gaussian $1D$. The $0D$ and $1D$ come from $nDmD$, where $n$ and $m$ are the dimensionalities of the measurement domain and the dependent variable, respectively (PATAKY; VANRENTERGHEM; ROBINSON, 2016).

In $nDmD$ data sets, the physical nature of the variables changes in the $m$ components, but not in the $nD$ measurement domain. Throughout the study, $0D$ and $1D$ represent $0D1D$ and $1DmD$, respectively. The study focused on only one statistical question: the probability of false positives in a single two-sample $0D1D$ t-test performed on $1DmD$ data. The two-sample t-test was applied to demonstrate the magnitude of the false positive problem and because the problem is exacerbated in more complex designs like ANOVA. The authors also recognize that many other issues should be considered when performing statistical analyses, including sample size, non-sphericity, normality, outliers, etc (PATAKY; VANRENTERGHEM; ROBINSON, 2016).

Just as it is useful to consider each of these questions individually, the authors also state that it is equally useful to consider the $0D$ vs. $1D$ individually, because this problem is relevant to all $1D$ data analysis, but it had not been explicitly addressed in the literature. The authors also highlighted that there are some custom open-source packages specifically for $1D$ analysis, but the available options are still in relatively early stages of development. Thus, simple interfaces to $1D$ procedures in commercial packages would remove a major barrier to accessibility of $1D$ procedures in Biomechanics (PATAKY; VANRENTERGHEM; ROBINSON, 2016).

A study on the influence of backpack carriage on gait initiation (GI), center of pressure (COP), and GRF time-series in three conditions was analyzed to capture features of the entire COP and GRF time series, rather than a few discrete variables. A vector analysis was conducted using SPM methods, which were inserted into a custom-written MATLAB program, to compute the resultant COP, process the data, and construct the figures. The authors did find different results using SPM and the traditional approach, besides other limitations of the experimental design, but they showed that SPM may improve analysis efficacy (VIEIRA; LEHNEN; NOLL; RODRIGUES *et al.*, 2016).

In 2017, a study to investigate the dynamics of GI on inclined surfaces evaluated the behavior of the COP and the center of mass (COM), based on the hypothesis that compared to the horizontal condition (HOR), the COP and COM excursion is higher in the upward (UP) condition and lower in the downward (DOWN) condition showing increased/decreased COM acceleration, and COP-COM vector to propel the body forward and upward/downward against/in favor of gravity, respectively. To capture the features of the entire COP and COM time series rather than discrete variables, the authors applied a vector analysis using SPM methods and concluded that compared to the HOR condition, both UP and DOWN conditions significantly alter COP and COM during all GI phases. The authors also presented the limitations, such as how SPM results may be affected by normalization of COP. The authors conclude that further studies may be necessary to overcome these problems and confirm the validity of the results (VIEIRA; DE BRITO; LEHNEN; RODRIGUES, 2017).

Furthermore, in 2017 GRFs between forward (FW) and backward (BW) walk was compared to test the hypothesis that the FW and BW kinematic patterns would be accompanied by the corresponding symmetrical kinetic behavior. In contrast to previous studies based on discrete GRF values, they captured data from the entire GRF time series and conducted a vector analysis using SPM methods. Looking over the entire time series, SPM can reveal significant differences in parts of the support phase that may be of special interest. Also, they compared the mean EMG profile between FW and BW during the stance phase, concluded that there is asymmetric kinetic behavior and different patterns of muscle activation throughout the stance phase between FW and BW, and that discrete value analyzes are not allowed for the identification of differences between FW and BW at different time intervals than those in which discrete values were identified (MAHAKI; SOUZA; MIMAR; VIEIRA, 2017).

The latest work reviewed that used SPM in a biomechanical setting, with the purpose of identifying secondary deviations in lower and upper body kinematics, as well as lower body joint moments caused by a unilateral knee brace-induced knee flexion contracture (KFC), in a sample of young healthy adults during gait. The knee brace was thereby locked at 30° and 60° of flexion to increase the chances for simulating an actual KFC of 30°. SPM was applied to compare continuous data instead of pre-defined parameters and several identified group differences were evaluated for their clinical relevance (SOTELO; EICHELBERGER; FURRER; BAUR *et al.*, 2018).

Appendix A presents the main findings as a chart summary and it complements this narrative resume of the state of the art of SPM applied to GRFs data analysis.

Clearly, there is opportunity to increase the number of studies that focus on SPM application on GRF experimental data, so we aim to contribute to the expansion of the use of the SPM method for studies in biomechanics, discussing methodological issues, challenges, limitations and advantages, as well as implications of the results and potential benefits of its applications.

# 3. THEORETICAL FOUNDATIONS

## 3.1. Human Locomotion: Static and Dynamic Aspects

### 3.1.1. Forces and Newton Laws of Motion

Forces can start, interrupt, or change the motion of the bodies on which they act, and can change its shape by deformation. Forces are always applied by one body on another body. A push on an object uses a muscular effort to produce a movement that has the direction of this push. A pull on an object on the opposite side will change the direction of the motion. Force is measured in newtons (N) in the International System of Units (SI). Three types of forces: gravitational force, muscle force, and friction will be discussed to understand forces acting during gait (OKUNO; FRATIN, 2014).

Force is an example of a vector quantity, and it is indicated by $\vec{F}$ or by boldface letter $\boldsymbol{F}$. Vectors are characterized by both magnitude and direction and can be represented graphically or mathematically. In a diagram, a vector is represented by an arrow whose direction determines the line of action; its length obeys a scale and is proportional to the magnitude or intensity of the force. A system of coordinates can be used to represent a force vector. In the case of rectangular coordinates, a force can be described through its projection on each axis. The sign of a rectangular component is positive (+) or negative (−) when the arrow is directed upward and to the right or down and the left, respectively. When two or more forces act on a body, it is possible to determine a force called resultant force, which can produce the same effect as all forces acting together (OKUNO; FRATIN, 2014).

*Newton's First Law of Motion (Law of Inertia):* it refers to the fact that a body will only change its state of motion (rest or uniform motion) if it experiences a net external force (resultant force). In other words, there are two equilibrium situations which rise from this law, static and dynamic. On static equilibrium both acceleration and speed are equal to zero, while on dynamic equilibrium the velocity is constant which implies that the acceleration is zero (OKUNO; FRATIN, 2014).

*Newton's Second Law (Mass and Acceleration):* When a nonzero resultant force acts on a body, it causes a change in the vector velocity, i.e., an acceleration $\vec{a}$. There is an inverse relationship between this acceleration and the mass $m$ of the body, and a direct relationship with the net force $\vec{F}$, that is,

$$\vec{a} = \vec{F}/m$$

Then, we can write that $\vec{F} = m\vec{a}$. The unit of velocity in SI is $m/s$, and as acceleration is given by the rate of change of speed $\Delta\vec{v}$ with time Δt, its unit in SI is $m/s^2$. Therefore, the unit of force is $kgm/s^2$ which receives the name newton, $N$, because of Isaac Newton (1642–1727) (OKUNO; FRATIN, 2014).

Newton's Third Law (Action and Reaction) states that a force is the result of two bodies interacting. The third law says that for every action force, there is an equal-intensity but opposite-direction reaction force. These action and reaction forces act on different bodies. During the contact phase of gait, reaction forces are exerted by the ground on the foot (OKUNO; FRATIN, 2014).

### 3.1.2. Some Specific Forces and Quantities Related to Forces

*Weight:* is the force exerted by the earth on bodies attracted to it (not always in contact); it is also known as gravitational force. This force is always directed toward the center of the earth. The intensity of the weight vector $W$ is thus defined by: $F = W = mg$, where $m$ is the body's mass in kilograms (kg) and $g$ is the acceleration of gravity, which is about $9.8\ m/s^2$ on the earth's surface. The force that the body puts on the earth (reaction force R) and acts on the center of the earth in reaction to the weight $W$ (action) exerted by the earth on a body. Its magnitude is the same as the weight and is in the opposite direction (OKUNO; FRATIN, 2014).

*Weight and Center of Gravity:* The resultant weight of a large body made up of small pieces will be equal to the total of the gravitational forces acting on each of these little components. The center of gravity (CG) is the place at which the resulting weight is applied, and it is the point in which all of the body's mass is concentrated. The CG is at the geometric center of homogeneous regular form objects (OKUNO; FRATIN, 2014).

An imaginary vertical line crossing through the CG must pass through the area specified by the support points for a body to be in rotational equilibrium. In the case of a person standing erect with both feet equally sustained by the ground, the area that delimits the points of support involves both feet, as illustrated in Figure 2. The larger is the support area the greater is the stability. It is why it is harder to remain in equilibrium standing on the tiptoe. Also, the center of gravity position may be located outside the body because it is a function of the body mass distribution (OKUNO; FRATIN, 2014).

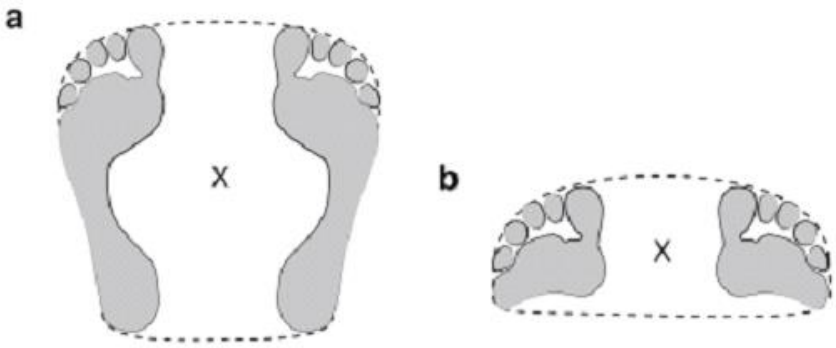



*Figure 2 Projection on the ground of the center of gravity (x) of a person standing erect: (a) over the area of support of the feet, and (b) over the area of support on the tip of the feet (OKUNO; FRATIN, 2014).*

*Muscle Forces:* Muscles produce forces that control posture and movement and are responsible for all body actions; the human body contains roughly 600 muscles whose primary function is muscle contraction, these are made up of many fibers, each of which contains cells that can contract when activated by nerve impulses from the brain. Usually, tendons attach muscles to two different bones. The cross-sectional area of a muscle dictates the maximum force that a muscle can exert and is inherent to the structure of muscle filaments It does not depend on the size of the animal (OKUNO; FRATIN, 2014).

*Contact Force:* Consider an object at rest on a surface. The object experiences a force $\overrightarrow{\mathrm{W}}$ due to the gravitational pull of the earth. As the object is at rest, the resultant of all applied forces must be zero. Therefore, another force of equal magnitude and on the opposite direction is applied by the surface of contact. This is the contact force or normal force $\vec{\mathrm{N}}$, which is perpendicular to the surface. The reaction to the force $\overrightarrow{\mathrm{W}}$ is exerted on the earth, and the reaction to N is another contact force $\vec{\mathrm{N}}^{*} = -\vec{\mathrm{N}}$, exerted by the block on the surface of contact (OKUNO; FRATIN, 2014).

*Forces of Friction* $(\overrightarrow{f})$*:* it is a force which a surface in contact with the body applies on it, when submitted to a force and it is parallel to the surface of contact. Friction forces appear in bodies in motion or on the verge of moving. It has an opposing direction to that of the externally applied force and it opposes the movement, and it is originated from the roughness of both surfaces in contact. When we walk or run, as the heel of the foot touches the ground, the foot pushes it in a forward direction and the ground exerts a frictional force in the backward direction, preventing the person from slipping. When the toe leaves the ground, the frictional force prevents the toe from

slipping backward. Therefore, we would not be able to walk or run on a surface without friction (OKUNO; FRATIN, 2014).

*Pressure:* the concept of pressure is associated with the force applied to a body. Pressure $p$ is defined as the force per unit area exerted perpendicularly on a surface. The pressure is inversely proportional to the area where it is applied. The unit of pressure in SI units is $N/m^2$. The pressure exerted by the weight of a person will become larger as the contact area of this person with the ground becomes smaller (OKUNO; FRATIN, 2014).

*Center of Pressure (COP):* it is the location of the origin GRF, and during normal gait, the COP progresses from the heel to toe during the stance phase. Initially the COP is in a medial position at the heel and rapidly moves to the lateral side of the mid-foot. It remains there during the mid-support period and then rapidly transfers to the medial side of the fore foot. To maintain balance while walking, the support force has to be on the same line from the point of ground support (COP) to the COM (CHOCKALINGAM; HEALY; NEEDHAM, 2018).

*Torques*: torque or moment of a force, $M_F$, is a physical quantity associated with the tendency of a force to produce rotation about an axis. Torque is calculated by the product of the magnitude of force by the distance ($d_\perp$) from the line of action of force $\vec{F}$ to the axis of rotation. The distance $d_\perp$ is called the moment arm or the lever arm of the force $\vec{F}$. The segment that defines the lever arm is perpendicular to the line of action of the force and passes through the axis of rotation. The magnitude of the torque, $M_F$, is defined by: $M_F = \vec{F} d_\perp$, and its unit in the International System of Units (SI) is $N\,m$ (OKUNO; FRATIN, 2014).

*Moment of Inertia of the Human Body:* it expresses the difficulty of changing the state of movement of a body in rotation. In a human body, due to subjects' variation and body parts having irregular shapes, it is not a trivial task to mathematically determine the moment of inertia. To define the moment of inertia, it is necessary to specify first the axis of rotation. The human body rotates, when it is free of support, about three axes, called the principal axes. These are mutually perpendicular lines that pass through the center of gravity corresponding to the posture of the body. These axes are called transverse (vertical), anteroposterior, and longitudinal (medial-lateral) (OKUNO; FRATIN, 2014). Studying and modifying it, e.g. by changing body posture, may lead to improvement on sports performance.

*Angular momentum and Its conservation:* an object rotating about an axis has an angular momentum $L$ given by the product of the moment of inertia $I$ of the body and the angular velocity $\omega$. The angular momentum is conserved as the net total torque due to the external actions on the body is zero. As the value of $L$ is maintained constant, if $I$ increases, $\omega$ must decrease and vice versa (OKUNO; FRATIN, 2014).

*Angular impulse:* the product of external net torque and the duration of the action defines the angular impulse and is responsible for the variation of angular momentum of a body. The rotational motions of the human body occur about the axis that passes through its center of gravity. In this case, an important observation is that the weight force acting on the body's center of gravity does not produce torque and, hence, does not change its angular momentum. Torques are originated by impulse forces that will introduce or change the angular momentum of the body. In case that the impulse force is not applied, the body will maintain its state of rotation, that is, its angular momentum is conserved (OKUNO; FRATIN, 2014).

*Simple machines* detailed analysis helps us to understand the abilities of humankind. Levers, pulleys and inclined planes are examples of simple machines. Machines can be defined as mechanisms projected to perform specific tasks, to facilitate or enable human action (OKUNO; FRATIN, 2014).

*Work Done by a Force:* when an object moves caused by a force, this force performs work, so work is always done by a force, but the simple fact of a force being applied on a body does not imply that work is performed, because there must be a displacement in the direction of this force. The importance of work is the association with energy, so a body has an amount of energy related to the capacity to do work (OKUNO; FRATIN, 2014).

*Human Locomotion Equipment:* constituted of around two hundred bones, articulations, and muscles. Aside from providing structural support and containing and building soft tissues like muscle, fat, and skin, it also allows the human being to move and displace with substantial effort, i.e. do mechanical work. The interior structure and form of bones combine lightness and strength. Articulations that are well-lubricated slide gently and without friction, allowing for a wide range of movements (OKUNO; FRATIN, 2014).

*Articulations and Joints:* Joints are points where bones connect. Some are immovable, securely attaching the bones (joints of the skull). Articulations are moveable joints that allow you to move around. The different shapes of joints allow

different types of movement, such as rotation (shoulder) and plane motion (knee/elbow). These articulated bones have their extremities covered by soft cartilage which are filled with a viscous fluid that guarantees good lubrication. The structure formed by these bones and articulations are kept in place by the tough strands of its muscles and ligaments, which are also responsible for the structure movement within certain limits. Exceeding these ligament strength limits cause considerable damage (OKUNO; FRATIN, 2014).

*Muscle and Levers:* the skeletal muscles are attached to bones and generate movements. Voluntary control can make these muscles to contract or relax. An active force is the muscle force on the bone segment of the lever. A resistance force corresponds to the weight of the segment plus the weight of the external loads added to it. A great number of muscles work in pairs to produce a given motion (OKUNO; FRATIN, 2014).

*Levers:* a long rod, which, by the action of forces, may rotate about an axis or a pivot called fulcrum. In the movements performed by human bodies and studied by biomechanics, mechanical levers are present. The articulation corresponds to the axis of rotation (fulcrum or pivot), bones work as a rigid structure (rod) on which forces are exerted, and muscles and ligaments provide the force to move the loads. There are three forces present on a lever system: The action force, $\boldsymbol{F}_A$, or applied force is the force exerted by the muscle; The resistance force, $\boldsymbol{F}_R$, is the load; and the reaction force at the fulcrum. These forces are represented by vectors whose lines of action depart from the point of muscle attachment into the bone. The vectors follow the muscle direction, but not necessarily the anatomical direction of the entire muscle (OKUNO; FRATIN, 2014).

The levers are classified into three categories: (i) First-class lever: In the body, this system of levers is often used to maintain posture or balance. (ii) Second-class levers: these systems provide mechanical advantage and is rarely found in the human body. (iii) Third-class levers are predominant in the human body and are designed for increasing the speed of motion rather than increasing load capabilities. Figure 3 presents an example of each lever type (OKUNO; FRATIN, 2014).

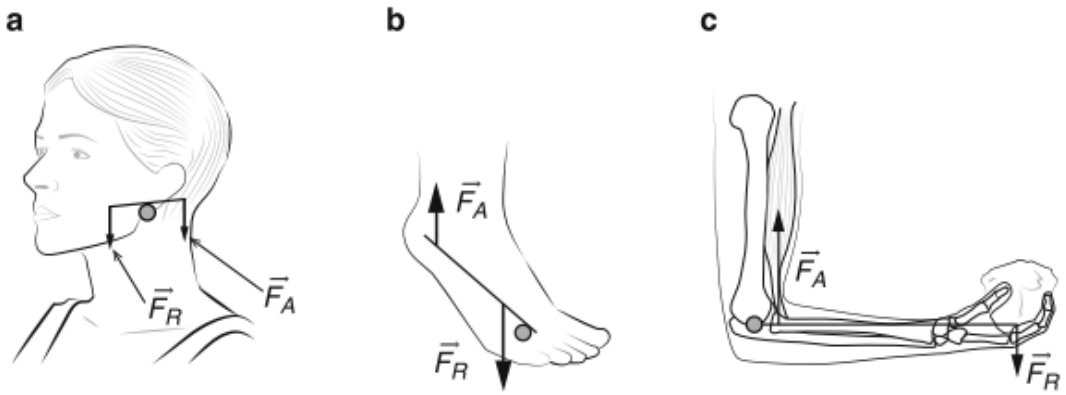



*Figure 3 Examples of levers in the human body. Levers of first, second and third class are represented in (a), (b), and (c), respectively (OKUNO; FRATIN, 2014).*

### 3.1.3. Walking Gait Cycle (GC)

A complete gait cycle is called a stride, it is the interval between two consecutive initial floor contact of the same limb characterizes a stride. In the middle of a stride the opposite foot contacts the surface to begins its next stance period. Thus the step is the interval between initial contacts of each foot, as shown in Figure 4 (PERRY; DAVIDS, 1992).

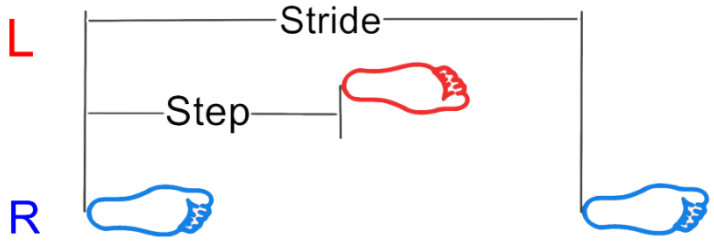



*Figure 4 Stride and step graphical representation for right and left foot. Modified from (PERRY; DAVIDS, 1992).*

The periodicity of leg movement is essential to the cyclic nature of the human gait. There are two basic requisites for the act of walking. First, there is the periodic movement of each foot from one position of support to the next. Secondly, applied through the feet there are sufficient GRFs to support the body as a whole system (PERRY; DAVIDS, 1992).

Conventionally, the beginning of the cycle is referred to as 0%. There is the assumption that successive walking gait cycles are equal which leads to normally describing walking as a single confined cycle. Even though this assumption is just ideal, it is a useful approximation (PERRY; DAVIDS, 1992).

The gait cycle is divided into a few phases to describe the processes that occurs during walking. The usual division of a cycle for a specific limb is into stance and swing phases. Stance refers to the period when the foot is in contact with the floor, and the swing phase when the foot is in preparation for the next foot strike. In healthy walking at a comfortable speed, the stance phase corresponds to 60% of the gait cycle (some references suggest 62% as a closer estimate), and the swing phase as 40% (or 38 %) of the gait cycle (BAKER, 2013; PERRY; DAVIDS, 1992; VAUGHAN, C. L.; DAVIS, B.; O'CONNOR, J. C., 1992).

Sub phases of gait: to provide the basic functions required for walking each stride presents an ever-changing alignment between the body and the support foot during stance and the selective advancement of the limb segments during swing. It results in a series of motion patterns performed by the hip, knee and ankle (PERRY; DAVIDS, 1992). Although the nomenclature is usually presented for one foot at a time, the same terminology applies to both feet, which for a normal walking pattern is half cycle behind of the opposite side, this way first double support for the right side is second double support for the left side, and vice versa (VAUGHAN, C. L.; DAVIS, B. L.; O'CONNOR, J. C., 1992). In normal gait there is a natural symmetry between the left and right sides, but in pathological gait an asymmetrical pattern very often exists. In the traditional nomenclature, the stance phase events are described as shown in Table 1 (VAUGHAN, C. L.; DAVIS, B. L.; O'CONNOR, J. C., 1992).

*Table 1 – Traditional stride events nomenclature by Vaughan (1992)*

| **Stance phase events** |
|---|
| 1. Heel strike initiates the gait cycle and represents the point at which the body center of gravity is at its lowest position |
| 2. Foot flat is the time when the plantar surface of the foot touches the ground. |
| 3. Mid stance occurs when the swinging (contralateral) foot passes the stance foot and the body’s center of gravity is at its highest position. |
| 4. Heel-off occurs as the heel loses contact with the ground and push-off is initiated via triceps *surae* muscles, which plantar flex the ankle. |
| 5. Toe-off terminates the stance phase as the foot leaves the ground |
| **Swing phase events** |

6. Acceleration begins as soon as the foot leaves the ground, and the subject activates the hip flexor muscles to accelerate the leg forward.
7. Mid swing occurs when the foot passes directly beneath the body, coincidental with midstance for the opposite foot.
8. Deceleration describes the action of the muscles as they slow the leg and stabilize the foot in preparation for the next heel strike.

The traditional nomenclature best describes the gait of normal subjects. However, there are several patients with pathologies, whose gait cannot be described using this approach. An alternative nomenclature with eight events sufficiently general to be applied to any type of gait and are shown in Figure 5 (VAUGHAN, C. L.; DAVIS, B. L.; O'CONNOR, J. C., 1992).

The analysis of the walking by phases identifies more directly the functional significance of the different motions occurring at the individual joints, and that it also provides a means for correlation of the simultaneous actions of the individual joints into patterns of total limb function. Each of the eight gait phases has a functional objective and a critical pattern of selective synergistic motion to accomplish the goal of identifying differences in pathological gait. She separates the sequential combination of the phases into three basic tasks: weight acceptance (WA), single limb support (SLS) and limb advancement (LA) (PERRY; DAVIDS, 1992).

There are several problems with Perry's subdivision, and there is a simpler division of both single support and swing into three subphases of equal duration: early, middle and late stance/swing phases (BAKER, 2013). Table 2 presents a summary of traditional gait phases based on Vaughan and Perry developments. Table 3 presents the primary muscles activated during gait of healthy subjects extracted from (MALANGA; DELISA, 1998).

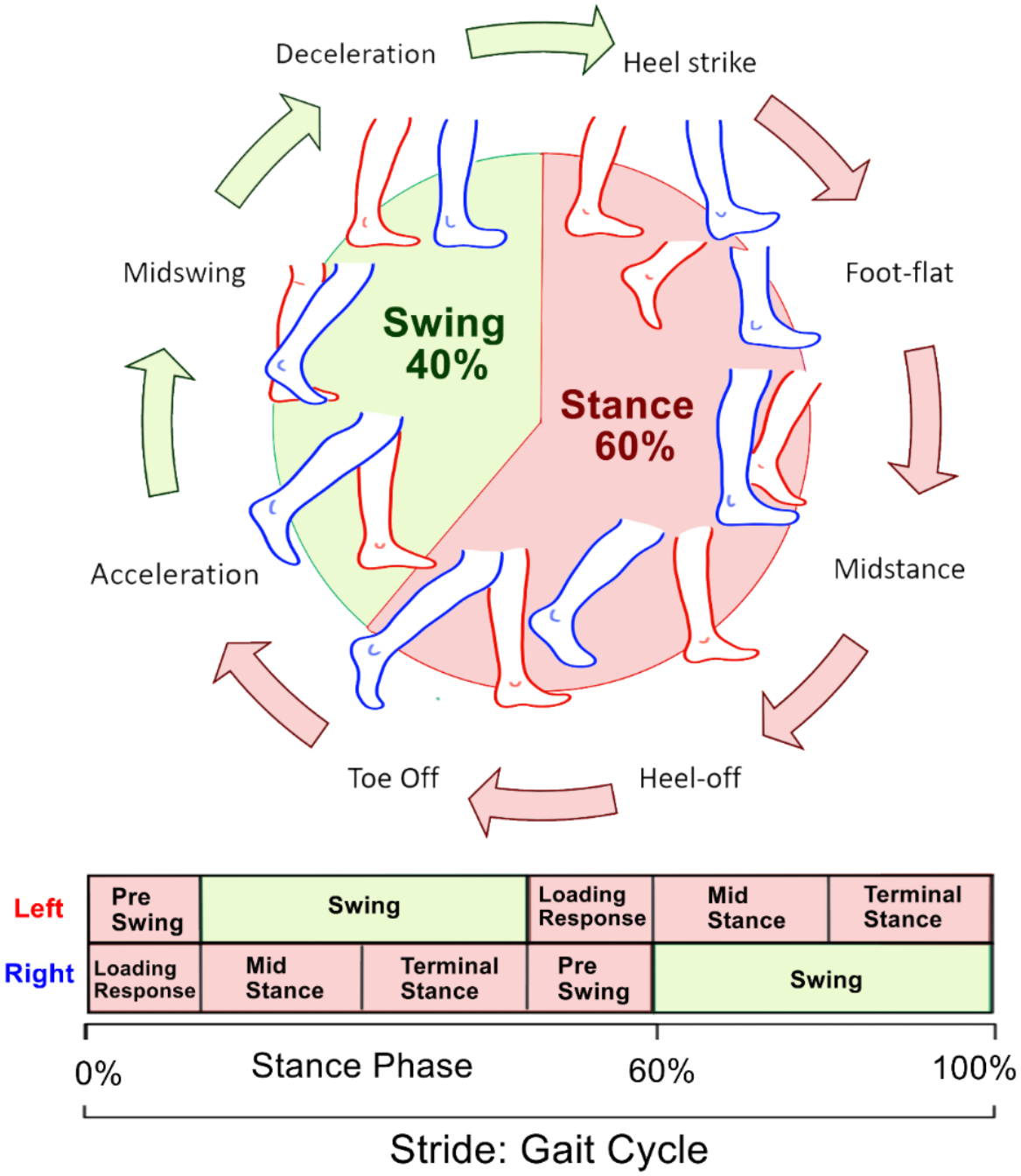



*Figure 5 Overview of the phases and subphases in the walking cycle. Modified from (PERRY; DAVIDS, 1992; VAUGHAN, C. L.; DAVIS, B. L.; O'CONNOR, J. C., 1992).*

*Table 2 Gait Phases description. Adapted from (PERRY; DAVIDS, 1992)*

| **1 GAIT CYCLE(GC) = 1 STRIDE = STANCE (60%GC) + SWING (40%GC)** | | **Interval (%GC)** | **Description** | **Objectives** |
|---|---|---|---|---|
| **TASK A- Weight Acceptance:** Three functional patterns are needed in this interval: shock absorption, initial limb stability and the preservation of progression. There is the abrupt transfer of body weight onto a limb that has just finished swinging forward and has an instable alignment. | Phase 1 - Initial Contact | 0-2 | This phase includes the moment when the foot just touches the floor. The joint posture determines the limb's loading response pattern. | The limb is positioned to start stance with a heel rocker. |
| | Phase 2 - Loading Response | 0-10 | This is the initial double stance period. The phase begins with initial floor contact and continues until the other foot is lifted for swing. | To absorb shock, weight-bearing stability and to preserve progression. |
| **TASK B - Single Limb Support:** During this interval, one limb has the total responsibility for supporting bodyweight in both the sagittal and coronal planes while progression must be continued. | Phase 3 - Mid Stance | 10-30 | It is the first half of the single limb support interval. It begins as the other foot is lifted and continues until body weight is aligned over the forefoot. | To progress over the stationary foot limb and for trunk stability. |
| | Phase 4 - Terminal Stance | 30-50 | It completes single limb support. It begins with heel rise and continues until the other foot strikes the ground. Throughout this phase body weight moves ahead of the forefoot. | For progression of the body beyond the supporting foot. |
| **TASK C - Limb Advancement:** To meet the high demands of advancing the limb, preparatory posturing begins in stance. Then the limb swings through three postures as it lifts itself, advances and prepares for the next stance interval. | Phase 5 - Pre-Swing | 50-60 | It begins with initial contact of the opposite limb and ends with ipsilateral toe-off. While the abrupt transfer of body weight promptly unloads the limb, the unloaded limb uses its freedom to prepare for the rapid demands of swing. | To position the limb for swing. |
| | Phase 6- Initial Swing | 60-73 | It begins with lift of the foot from the floor and ends when the swinging foot is the opposite stance foot. | For foot clearance of the floor and for advancement of the limb from its trailing position. |
| | Phase 7 - Mid Swing Phase | 73-87 | It begins as the swinging limb is opposite to the stance limb. The phase ends then the swinging limb is forward, and the tibia is vertical (i.e. hip and knee flexion postures are equal). | For limb advancement and foot clearance from the floor. |
| | Phase 8 - Terminal Swing | 87-100 | It begins with a vertical tibia and ends when the foot strikes the floor. Limb advancement is completed as the leg moves ahead of the thigh. | To complete limb advancement and to prepare the limb for stance. |

*Table 3 Primary muscular activity during the gait cycle*

| Muscular Activity | Muscles | Period |
|---|---|---|
| **Shock absorbers** | Quadriceps and dorsiflexors | Weight loading<br>(Weight acceptance) |
| **Stabilizers** | Gluteous Maximus, Medius, & Minimus;<br>Tensor Fascia Lata;<br>Erector Spinae | Stance Phase |
| **Foot Lift Off** | Flexor Digitorium Longus<br>Flexor Hallucis Longus<br>Peroneus Longus and Brevis<br>Tibialis Posterior | Weight unloading |
| **Accelerators** | Adductor Longus<br>and Magnus Iliopsoas Sartorius | Weight-Unloading |
| **Foot Controllers** | Extensor Digitorum Longus<br>Extensor Hallucis Longus<br>Tibialis Anterior | Swing-Phase |
| **Decelerators** | Gracilis<br>Semimembranosus<br>Semitendinosus<br>Biceps Femoris | Swing-Phase<br>mid-swim; to initial-contact |

### 3.1.4. Ground Reaction Force during standing and walking

*System of Parallel Forces:* when forces acting on a rigid object are applied perpendicularly to a given segment of a straight line it is identified as a system of parallel forces, as on standing position. The algebraic equations to obtain unknowns by calculation are provided by the conditions for static equilibrium for the forces and their respective torques. As an example, we can consider the magnitude of the contact forces exerted by the ground on the right foot $N_R$ and on the left foot $N_L$of a person standing erect, as shown in Figure 6 (OKUNO; FRATIN, 2014).

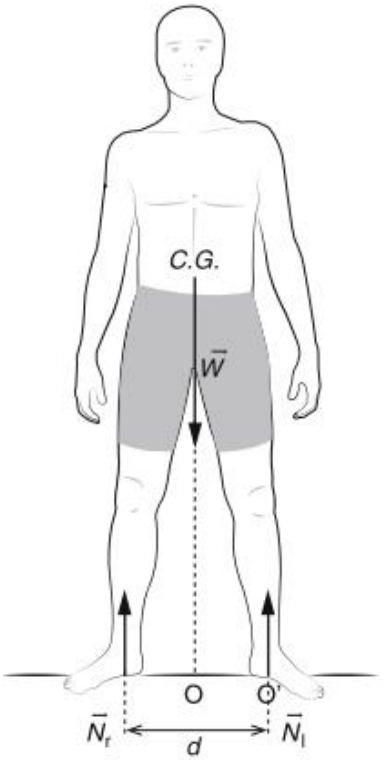



*Figure 6 Example of a parallel force system acting on support area of feet during the standing position. Extracted from (OKUNO; FRATIN, 2014)*

The mass of this person is 70 kg and the position of the center of gravity (CG) is indicated in the figure. The feet are 30 cm apart from each other and the line that passes through CG passes midway between the feet. First, we transport the weight force $W$ to the level of the ground and choose $O'$ as the axis of rotation, at the left foot, for convenience, and apply the second condition of equilibrium:

$$M_T = (700\ N)(0.15\ m) - Nr(0.30\ m) = 0,$$

$$N_R = \frac{105\ N\ m}{0.30\ m}\ 350\ N.$$

Then by using the second condition of equilibrium and assuming a gravity acceleration equal to $10\ m/s^2$ we obtain:

$$W\ = N_R\ +\ N_L\ =\ \ 700\ N,$$

so, as $N_R = \ 350\ N, N_L = \ 350\ N$. This equation also makes it possible to calculate body weight if $N_R$ and $N_L$ are known quantities.

*System of Nonparallel Forces:* the forces that occur inside or on the body are not always parallel but form an angle between them. To solve these problems, it is useful to decompose the forces in some convenient orthogonal directions. To do this, an interesting way to choose the x and y axes is that in which one of the components of the torque becomes zero, not causing rotation although it can cause compression or tension on an articulation (OKUNO; FRATIN, 2014). As we have seen on the parallel force systems section, in quiet standing, the ground reaction is constant, being equal in magnitude and opposite in direction to body weight. During normal gait the GRF varies above and below resting body weight according to vertical acceleration. The reaction force is not affected by constant speed up or down (KIRTLEY, 2006).

*Forces on the Hip:* during walking there is a brief period where we stand erect on one leg, and it changes at every footstep. At this point, the center of gravity lies on an imaginary line passing simultaneously through the vectors weight force $W$ of the body and normal force $N$, which equilibrates $W$. $N$ is exerted by the floor on the foot that touches the ground (OKUNO; FRATIN, 2014).

A contact force $C$ larger than twice the body weight force is exerted by the articulation of the hips on the head of the femur of the leg that supports the weight at each footstep. This contact force changes drastically, according to the body weight sustained by one foot or the other alternately during walking (OKUNO; FRATIN, 2014).

The muscles gluteus medius, gluteus minimus and tensor of fascia lata femoris exert the hip abductor muscle force $F$ on the great trochanter of the femur. Figure 7 shows a geometric model with the applied forces on the inferior right limb with the foot on the floor (OKUNO; FRATIN, 2014).

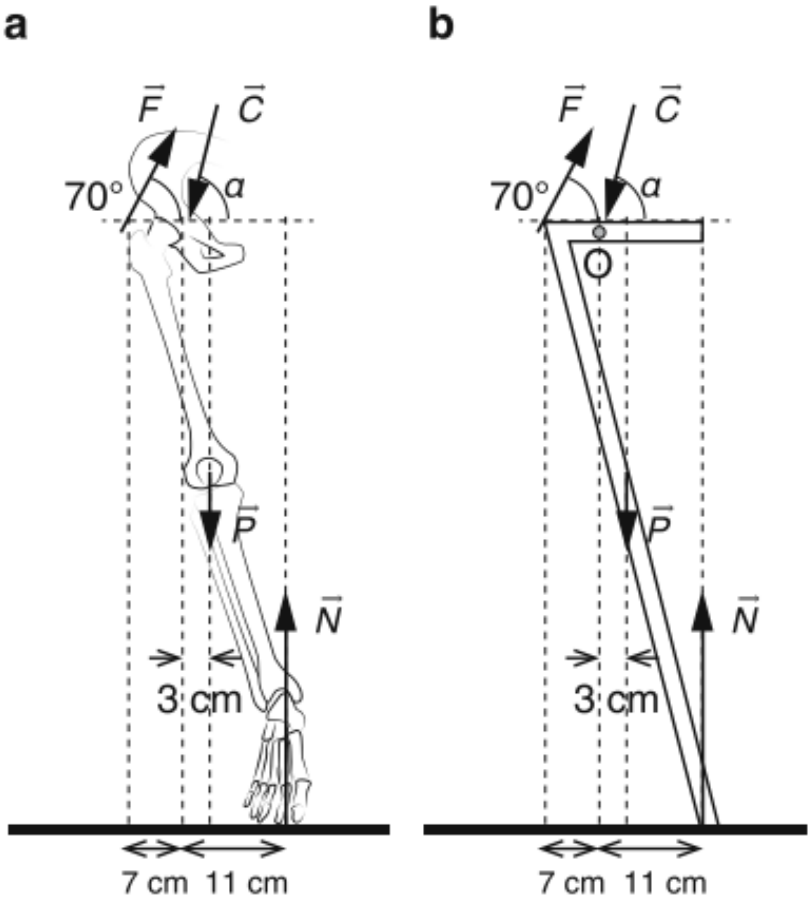



*Figure 7 (a) An adult standing on his right foot, in static equilibrium. P is the weight of the set thigh-leg-foot and N, the normal reaction to the weight force W of the body. (b) Sketch of the geometric model of the leg of (a) (OKUNO; FRATIN, 2014).*

The line of action of the weight force passes through the center of gravity placed on the foot which is on the floor during walking. The forces involved in walking (dynamic) can be analyzed considering that the body is in static equilibrium momentarily, while one of the feet is on the ground (OKUNO; FRATIN, 2014).

### 3.1.5. Force Platforms

There are different types of force platforms (or force plates) that are applied to record GRFs. Most of them either consist of strain gauge or piezoelectric crystal transducers. Both types of platform measure forces indirectly, based on a calibration matrix. There are also pressure measurement systems which are available both as in-shoe systems and platforms. These can give an indication of the vertical component of the GRF (CHOCKALINGAM; HEALY; NEEDHAM, 2018).

The GRF acting on the foot during upright movements is a major external force, normally recorded and measured in three dimensions using a force platform. It

comprises of a single vertical and a pair of horizontal shear force components acting in the anterior-posterior and medial-lateral directions. Force platform data also provide coordinates of the point of application of the resultant vector relative to the platform origin (CHOCKALINGAM; HEALY; NEEDHAM, 2018).

The force can be expressed in two ways: the force on the platform (action) or the reaction on the body (reaction). These two ways have equal magnitude and opposite directions, which means opposite signs. Figure 8 presents the common disposition of GRFs on a force plate. It is advised to label shear forces as $F_{AP}$ and $F_{ML}$ rather than $F_x$ and $F_z$, to avoid confusion (KIRTLEY, 2006).

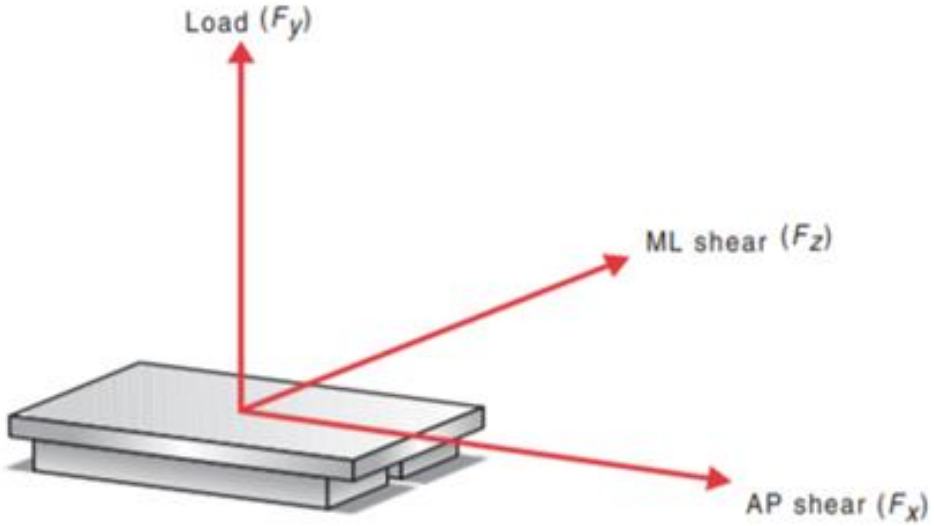



*Figure 8 Graphical representation of GRF components measured by a force platform. Source:Modified from (Kirtley, 2006)*

Major forces involved in human locomotion are the ground reaction force (GRF), joint reaction force (JRF), and muscle force (MF) (CHOCKALINGAM; HEALY; NEEDHAM, 2018). Forces act to accelerate a body and do so independent of where the force is applied. The most important force in gait analysis is the ground reaction, it is the force exerted by the ground on the foot (provided by a supporting horizontal surface) (BAKER, 2013; CHOCKALINGAM; HEALY; NEEDHAM, 2018).

It is called a reaction because it is the equal and opposite force to that which the combination of gravity and muscle activity within the body exerts upon the ground. Many gait analysis services do not plot the GRFs, but it is impossible to understand joint moments properly without knowledge of how it is acting (BAKER, 2013). It is usually visualized as an arrow, with the length of the arrow representing the magnitude of the force, and it can only be plotted on gait graphs if we describe it in terms of

individual components (BAKER, 2013). These forces can be resolved into three components, namely, vertical, medial-lateral, and anterior-posterior, which are defined in a global coordinate system and are shown in Figure 9 (BAKER, 2013; CHOCKALINGAM; HEALY; NEEDHAM, 2018).

The vertical component is always positive, while the other two may be positive or negative. The ground reaction supports the body against gravity and also accelerates the body's center of mass. The vertical component of the ground reaction is always much larger than the other components because resisting gravity takes much more force than accelerating it during walking and running (BAKER, 2013). The point of application of the resultant GRF is represented by the center of pressure (COP) (CHOCKALINGAM; HEALY; NEEDHAM, 2018).

The movement of the body segments originates from a force exerted by a muscle which is transmitted through its tendon to the bone. Muscle generates tensile force and is not able to generate compressive force, this way we can say that muscles can only pull and cannot push. Contact forces of ligaments and bones exerted across a joint are transmitted from one body segment to another. These forces are called joint reaction forces (JRF) and are essential in some aspects of musculoskeletal biomechanics, especially in the design of prosthesis and prosthetic joints (CHOCKALINGAM; HEALY; NEEDHAM, 2018). Gait analysis is important in orthopedics, physical medicine and rehabilitation and has been applied to the assessment of several clinical conditions (CHOCKALINGAM; HEALY; NEEDHAM, 2018). Figure 10 presents the GRF patterns and its peaks during stages of the gait cycle, which are usually detected for traditional GRF analysis..

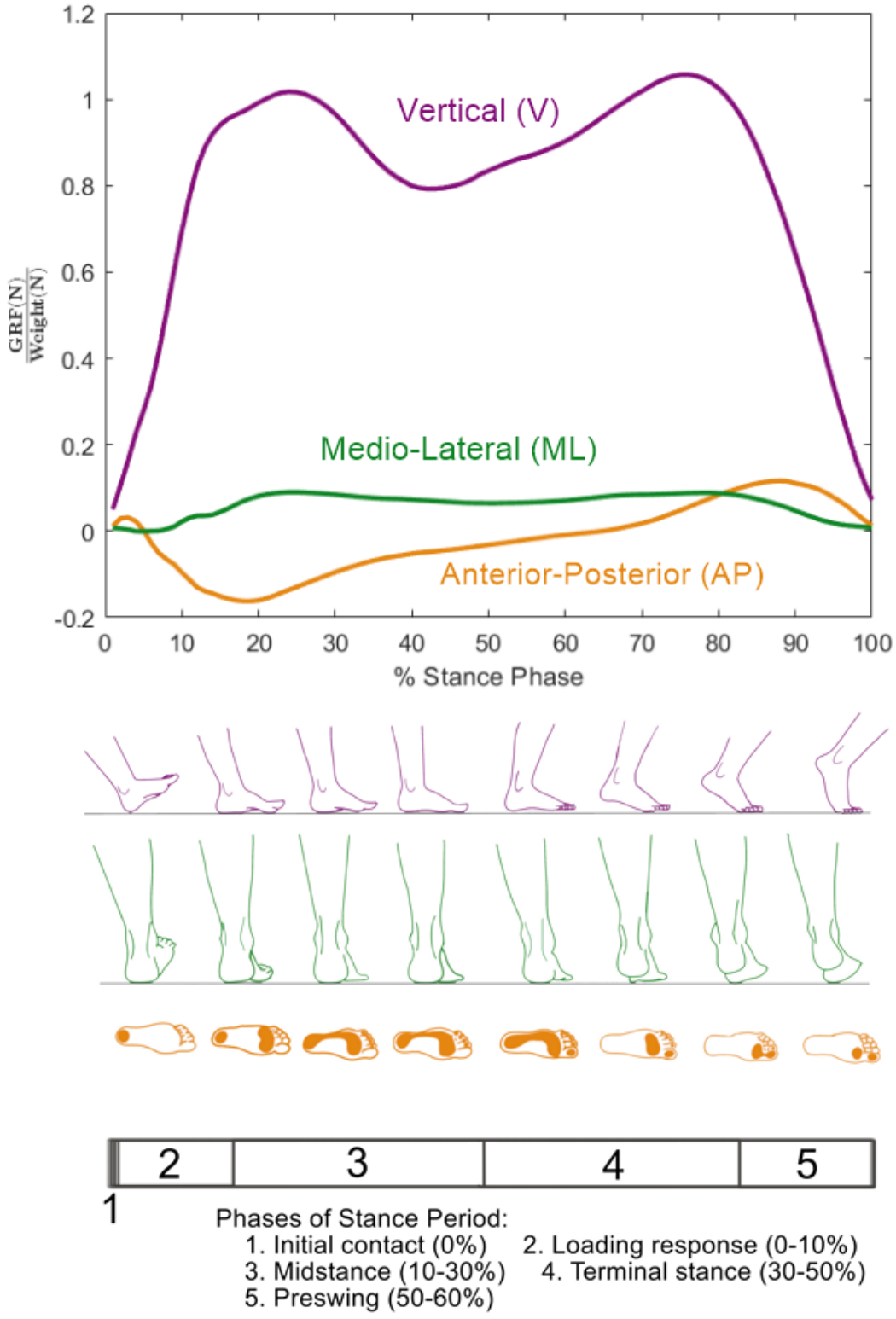
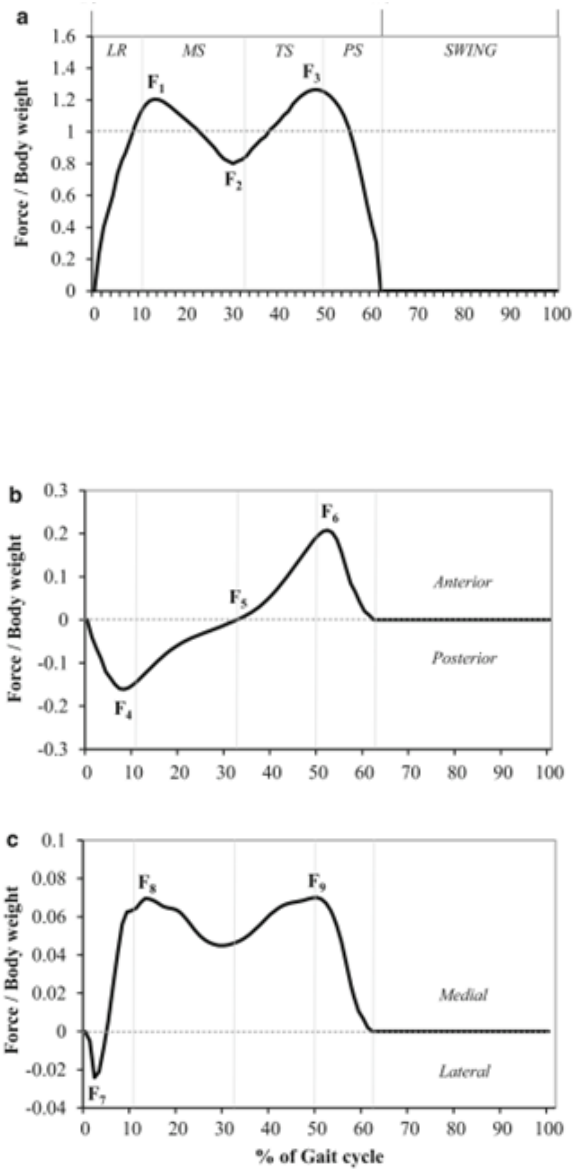



*Figure 9 GRF components and graphical representation of foot position during stance phase. (The Author, 2021) based on (PERRY; DAVIDS, 1992) and (NOREILS, 2014).*




**Vertical Component of the GRF:** At early stance (from initial contact throughout the loading response phase) the vertical GRF progressively increases as the weight of the individual is transferred from the back foot to the front foot and the body decelerates downward. During the initial period of mid-stance peak loading is reached due to weight acceptance and an increase in muscular forces as an individual transits from double to single limb support. The center of mass displaces in an upward direction through mid-stance as the knee extends (F1). The center of mass experiences a deceleration near its highest position which presents a reduction in the vertical GRF below the actual body weight (F2). The vertical GRF increases until a second peak (F3) is reached in terminal stance, which relates to the foot pushing against the floor as a result of the increase in the activity of the ankle plantar flexors and from the deceleration of the center of mass as the body weight falls forward (Chockalingam *et al.* , 2018).

**Anterior-Posterior Component of the GRF:** at initial contact the foot applies a force in an anterior direction causing the body to decelerate. A posterior shear force results from this presenting a peak during early stance. With the movement of the body over the stance limb, the posterior shear forces decrease until the end of mid-stance (F5), which is a crossover point. Then the center of mass moves ahead of the foot, and the foot applies a force in a posterior direction causing an anterior shear force that propels the body forward. The peak of anterior GRF (F6) occurs during late stance, followed by the reduction related to the transference of force to the front foot (Chockalingam *et al.* , 2018).

**Mediolateral Component of the GRF:** At initial contact there is a medially directed shear force applied by the foot to the ground. It creates a lateral shear force presenting a peak during loading response (F7). A force directed laterally is applied by the foot to the ground from mid load response which results in a medial shear force that is characterized by two peaks approximately at the beginning of mid-stance (F8) and toward the end of terminal stance (F9) (Chockalingam *et al.* , 2018).

*Figure 10 Vertical (a), Anterior-Posterior (b) and Medial-lateral (c) components of the GRF (Chockalingam et al., 2018).*

## 3.2. **Continuous GRF Analysis – SPM**

The design of continuous statistical processes to test hypothesis concerning regionally specific effects is known as statistical parametric mapping. Statistical parametric maps (SPMs) are images or fields with values distributed according to a known probability density function, commonly the Student's t or F-distributions, under the null hypothesis. SPM has come to refer to the application of the general linear model (GLM) and random field theory (RFT) theories in combination to analyze and form classical inferences about topological characteristics of statistical parametric maps (SPM) (PENNY; FRISTON; ASHBURNER; KIEBEL *et al.*, 2011).

The following mathematical explanations were extracted from the vast literature of the software package creator of SPM1D (www.spm1d.org), and explains, briefly, the mathematics under SPM method (PATAKY, 2010; PATAKY, 2012; PATAKY; ROBINSON; VANRENTERGHEM, 2013).

Our purpose is to apply the SPM method on GRF data, so we are not going to focus on all the mathematical details and the genesis of the theories. We are going to present the main aspects which were previously refined by Todd C. Pataky on his vast amount of work on the subject.

### 3.2.1. General Linear Model

Mass-univariate general linear (GLM) models can summarize the relationship between experimental observations $\mathbf{Y}$ and experimental design $\mathbf{X}$:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where β is an $(\boldsymbol{J} \times \boldsymbol{K})$ matrix of unknown regression parameters and ε is a matrix of residuals. $\boldsymbol{Y}$ and $\boldsymbol{\varepsilon}$ are $(\boldsymbol{I} \times \boldsymbol{K})$, $\boldsymbol{X}$ is $(\boldsymbol{I} \times \boldsymbol{J})$, where: $\boldsymbol{I}$ is the number of observations; $\boldsymbol{J}$ is the number of experimental factors and $\boldsymbol{K}$ are discrete measurement points (nodes). An experimental observation $\boldsymbol{Y}$ is an n-dimensional sampling of a scalar field that can be flattened into a K-vector. A complete experiment produces $\boldsymbol{I}$ flattened $\boldsymbol{K}$-vectors. The least-squares of $\boldsymbol{\beta}$ can be estimated as:

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{X}^{+}\boldsymbol{Y} = (\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{Y},$$

producing errors,

$$\boldsymbol{\varepsilon} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}$$

$\boldsymbol{X}^{+}$is the Moore-Penrose pseudo-inverse matrix of $\boldsymbol{X}$:

$$X^{+} = \frac{X^{T}}{(X^{T}X)}.$$

It is applied to deal with matrices that are not square, square relates to a matrix with even number of rows and columns. Like $Y$ (dataset), the fitting models $X\widehat{\beta}$ are $(I \times K)$. In general, a large proportion of variability can be explained using this approach. After estimating parameters $\widehat{\beta}$ and residuals $\varepsilon$, the next task is to compute test-statistic values.

The GLM equation offers arbitrary linear testing, the generalized t test will be considered for simplification. The first nodal variance estimate. Nodal variance $\widehat{\sigma}_k^2$ is estimated as the ratio:

$$\widehat{\sigma}_k^2 = \frac{(\varepsilon^T\varepsilon)_{kk}}{v} = \frac{(\varepsilon^T\varepsilon)_{kk}}{I - rank(X)}$$

$(\varepsilon^T\varepsilon)_{kk}$ is the $k$th diagonal element of the $(I \times K)$ error sum of squares matrix $(\varepsilon^T\varepsilon)$ and $v$ is the error degrees of freedom. The nodal t statistic can then be calculated as:

$$t_k = \frac{\mathbf{c}^T\widehat{\beta}_k}{\widehat{\sigma}_k\sqrt{\mathbf{c}^T(X^TX)^{-1}\mathbf{c}}}$$

$\mathbf{c}$ is a $(J \times 1)$ contrast vector. The nodal values $t_k$ form a K-vector that can be reshaped into the original $nD$ sampling space and viewed in the context of the original data. The equation $t_k$is known as a statistical map and is referred to in the literature as $SPM\{t\}$. The contrast vector $\mathbf{c}$ assigns weights to the $J$ experimental factors and represents the experimental hypothesis.

As the above equations yield mean and variance curves, it is apparent that arbitrary linear statistical testing can be implemented with various choices of design matrices X and various combinations of the $\widehat{\beta}_k$ curves.

SPM procedures are conceptually identical to univariate procedures. Running a one-sample t test on 10 scalar values, for example, is almost identical to conducting a one-sample t test on ten vector fields. The differences are that SPM takes vector covariance into account when computing the test statistic, field smoothness and size into account when computing the critical test statistic threshold, and random field behavior into account when computing p values (PATAKY; ROBINSON; VANRENTERGHEM, 2013).

In a practical example, consider a set of curves representing independent experimental observations. After data preparation (registration), the data may be arranged into an $\boldsymbol{I} \times \boldsymbol{K}$ matrix $\boldsymbol{Y}$, where $\boldsymbol{I}$ and $\boldsymbol{K}$ are the numbers of curves and nodes per curve, respectively, and an experimental design can be represented by an $\boldsymbol{I} \times \boldsymbol{J}$ matrix $\boldsymbol{X}$, where $\boldsymbol{J}$ is the number of experimental factors. Given data $\boldsymbol{Y}$ and design $\boldsymbol{X}$, the entire family of linear parametric statistical tests can be implemented using a mass-univariate general linear model (GLM) (PATAKY, 2012).

### 3.2.2. SPM1D Independent Hotelling T² test

We applied SPM to test the null hypothesis by statistically examining the whole GRF time-series. All SPM analyses were implemented in MATLAB 2020a. In an I-component vector $\boldsymbol{y}$ which varies over $\boldsymbol{Q}$ points in space or time may be regarded as an ($\boldsymbol{I} \times \boldsymbol{Q}$) vector field response $\boldsymbol{y}(\boldsymbol{q})$. Given J responses, the mean vector field is:

$$\bar{\boldsymbol{y}}(\boldsymbol{q}) = \frac{\mathbf{1}}{\boldsymbol{J}} \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{J}} \boldsymbol{y_j}(\boldsymbol{q}).$$

The participants are represented by $\boldsymbol{j}$ indexes. The SPM's vector field analog to the independent t test is the independent Hotelling's T² test (PATAKY; ROBINSON; VANRENTERGHEM, 2013).

$$\boldsymbol{SPM}\{\boldsymbol{T^2}\} \equiv \boldsymbol{T^2}(\boldsymbol{q}) = \frac{\boldsymbol{J_1 J_2}}{\boldsymbol{J_1 + J_2}} (\bar{\boldsymbol{y}}_{\mathbf{1}}(\boldsymbol{q}) - \bar{\boldsymbol{y}}_{\mathbf{2}}(\boldsymbol{q}))^{\boldsymbol{T}} \boldsymbol{W}(\boldsymbol{q})^{-\mathbf{1}} (\bar{\boldsymbol{y}}_{\mathbf{1}}(\boldsymbol{q}) - \bar{\boldsymbol{y}}_{\mathbf{2}}(\boldsymbol{q}))$$

where subscripts 1 and 2 index the two groups being compared. Here **W** is the pooled covariance matrix and it represents the variances-within and correlations-between vector components across the J responses.

$$\boldsymbol{W}(\boldsymbol{q}) = \frac{\mathbf{1}}{\boldsymbol{J_1 + J_2 - 2}} \left( \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{J_1}} (\boldsymbol{y_{1j}} - \bar{\boldsymbol{y}}_{\mathbf{1}})(\boldsymbol{y_{1j}} - \bar{\boldsymbol{y}}_{\mathbf{1}})^{\boldsymbol{T}} \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{J_2}} (\boldsymbol{y_{2j}} - \bar{\boldsymbol{y}}_{\mathbf{2}})(\boldsymbol{y_{2j}} - \bar{\boldsymbol{y}}_{\mathbf{2}})^{\boldsymbol{T}} \right)$$

where the domain (q) is dropped for compactness. The GRF were analyzed as a three-component vector field $\boldsymbol{I} = \mathbf{3}$, $\boldsymbol{J} = \mathbf{12}$, $\boldsymbol{Q} = \mathbf{101}$, where $\boldsymbol{I}$, $\boldsymbol{J}$, and $\boldsymbol{Q}$ were the number of vector components, responses (participants) and time points (nodes) respectively. These solutions reduce to the univariate distribution (traditional scalar extraction) when the number of time points is equal to one ($\boldsymbol{Q} = \mathbf{1}$).

The notation $\boldsymbol{SPM}\{\boldsymbol{T^2}\}$ indicates that the test statistic $\boldsymbol{T^2}$ varies in continuous time (or space), forming a temporal (or spatial) statistical map. To clarify: “SPM” refers to the methodology, and $\boldsymbol{SPM}\{\boldsymbol{T^2}\}$ to a specific variable (PATAKY; ROBINSON; VANRENTERGHEM, 2013).

### 3.2.3. Random Field Theory – Statistical Inference

The following mathematical explanations were extracted from (PATAKY, 2010) supplementary material, and modified to explain 1D (time-series) data. Pataky refers to the work that were primarily developed for medical imaging application of SPM. The mathematics that supports topological statistical inference on an SPM is given by the Random Field Theory (RFT). Given the error degrees of freedom, the expected topological statistical inference on an SPM depends on field smoothness and search space geometry. Field smoothness can be estimated at each node by first computing normalized residuals $\boldsymbol{u}$.

$$\boldsymbol{u_j} = \frac{\boldsymbol{\varepsilon_j}}{\boldsymbol{\varepsilon_j^T \varepsilon_j}}$$

where $\boldsymbol{j}$ indexes the observations (responses), then assembling an ($\boldsymbol{J} \times \boldsymbol{n}$) gradient matrix for each node (1D), pixel (2D) or voxel (3D):

$$\dot{\boldsymbol{u}}_{\boldsymbol{q}} \equiv \begin{bmatrix} \boldsymbol{\nabla}(\boldsymbol{u_q})_{\boldsymbol{1}} \\ \vdots \\ \boldsymbol{\nabla}(\boldsymbol{u_q})_{\boldsymbol{J}} \end{bmatrix} = \begin{bmatrix} \frac{\boldsymbol{\partial}(\boldsymbol{u_q})_{\boldsymbol{1}}}{\boldsymbol{\partial_1}} & \cdots & \frac{\boldsymbol{\partial}(\boldsymbol{u_q})_{\boldsymbol{J}}}{\boldsymbol{\partial}} \\ \vdots & \ddots & \vdots \\ \frac{\boldsymbol{\partial}(\boldsymbol{u_q})_{\boldsymbol{J}}}{\boldsymbol{\partial_1}} & \cdots & \frac{\boldsymbol{\partial}(\boldsymbol{u_q})_{\boldsymbol{J}}}{\boldsymbol{\partial_n}} \end{bmatrix}$$

$\boldsymbol{\nabla}(\boldsymbol{u_q})_{\boldsymbol{j}}$ is the gradient of the $\boldsymbol{j}$th residual’s $\boldsymbol{q}$th node (time point), and $\frac{\boldsymbol{\partial}(\boldsymbol{u_q})_{\boldsymbol{i}}}{\boldsymbol{\partial_n}}$ is the $\boldsymbol{d}^{\mathbf{th}}$ component of that gradient vector. Nodal smoothness can be estimated as:

$$\widehat{\boldsymbol{W}}_{\boldsymbol{q}} = (\boldsymbol{4log2})^{\frac{\boldsymbol{1}}{\boldsymbol{2}}} |\dot{\boldsymbol{u}}_{\boldsymbol{q}}{}^{\boldsymbol{T}} \dot{\boldsymbol{u}}_{\boldsymbol{q}}|^{\frac{\boldsymbol{1}}{\boldsymbol{2}}} .$$

The full width at half-maximum (FWHM) of a Gaussian kernel that when convolved with uncorrelated Gaussian random field data produces the same smoothness as the normalized residuals $\boldsymbol{u_j}$ is estimated here with $\widehat{\boldsymbol{W}}_{\boldsymbol{q}}$. The expected size of suprathreshold $\boldsymbol{SPM}\{\mathrm{t}\}$ clusters increase as $\widehat{\boldsymbol{W}}_{\boldsymbol{q}}$increases, and it is a fact that RFT exploits. For simplicity it assumes isotropic smoothness. The shape of the $\boldsymbol{nD}$ space in which the data are located influences the expected topological properties of an $\boldsymbol{SPM}\{\boldsymbol{t}\}$.

The first step, assuming a $\boldsymbol{nD}$ dataset, is to assemble basic morphological characteristics of the $\boldsymbol{nD}$ space (density functions, count number of nodes, pixels and/or voxels), its global geometry can now be summarized by 'resel' or 'resolution element' counts. These resolution elements assume position independent smoothness $\widehat{\boldsymbol{W}} = \sum \widehat{\boldsymbol{W}}_{\boldsymbol{q}}/\boldsymbol{Q}$. Each resolution element is associated with an independent probability density function (Euler characteristic density) that directly depends only on the t threshold. These density functions can be used to compute a variety of topological expectations, like the number of suprathreshold nodes and clusters, for example. The final steps in RFT-based inference are thus to threshold an observed $\boldsymbol{SPM}\{\boldsymbol{t}\}$ at a suitably high value and then corroborate the observed topology with topological expectation, computing p-values for each cluster according, for example. The logic of RFT is that smooth random fields are expected to produce spatially broad suprathreshold clusters, but very broad and/or very high clusters are expected to occur with low probability. The key message is that a large suprathreshold cluster is the topological equivalent of a large univariate t value (PATAKY, 2010).

To conclude, for each suprathreshold cluster found, the RFT processes provide a single p value. These p values can be read as the likelihood that the observed suprathreshold cluster could have evolved from a smooth random process, given the field smoothness (calculated from) and the search space (a function of Q) (PATAKY, 2012). We are not going to dive into the mathematics of resolution elements, as our objective is to implement the method on our dataset (LAPAFI) and compare it with the external dataset (GaitRec).

#### 3.2.4. Post-hoc t test with Šidák p-value correction

Post-hoc analysis is done after the vector-field analysis (Hotelling's $T^2$) on each vector component separately (scalar fields), but only if statistical significance is reached. For scalar-fields the independent Hotelling's $T^2$ test becomes independent t-test (PATAKY; ROBINSON; VANRENTERGHEM, 2013). To maintain a family-wise error of α=0:05, Šidák thresholds of p=0.0170, were used to correct for the I=3 vector components of the datasets LAPAFI and GaitRec respectively.

## 3.3. **Characteristics of Walking Gait in Aged Adults**

### 3.3.1. Kinetics and Kinematics

Walking is learned, improved, and deprecated as our bodies age. This activity consists of the moving body being supported successively by one lower limb and the other (PRINCE; CORRIVEAU; HÉBERT; WINTER, 1997). Gait and mobility are different. Human gait refers to a walking style, to the locomotion achieved by using the lower limbs. During bipedal gait, the center of gravity is displaced forward. Mobility is the ability to displace in the environment with ease, and without restriction. The later does not necessarily uses the limbs to move the body from one place to another (CRUZ-JIMENEZ, 2017).

The dynamic control of upstanding stance is fundamental to perform daily living activities safely and efficiently (PRINCE; CORRIVEAU; HÉBERT; WINTER, 1997). Mobility disability in adults is an important factor for loss of independence. Equilibrium and locomotion are two fundamental components to the ability to walk. The capacity to assume an upright posture and maintain balance is called equilibrium. The ability to initiate and maintain rhythmic steps is called locomotion (CRUZ-JIMENEZ, 2017).

To maintain an upright posture, the erect spine rests on the sacral base conforming its natural curves to the center of gravity. To reach stability, both static and dynamic functions are balanced by the spine, including weight bearing and balance. It is accomplished by using structures like the anterior vertebrae, both anterior and posterior longitudinal ligaments, facets, and spinal muscles. Targeting energy conservation, the center of gravity is kept within the delimitation created by the feet in its base of support, and the static balanced spine posture benefits from GRFs to minimize muscle activation. These forces strategically cross the naturally conformed spine curves and the lower extremity joints, allowing balance and static control (CRUZ-JIMENEZ, 2017).

An unbalanced posture of the spine culminates in a three-dimensional motion of the spine, causing concurrent coupling of flexion, lateral flexion, and rotation of the muscles of the spine. There is neurologic control that improves the precision of the center of mass displacement, deviated from simple movements in the balanced posture like arm elevation, and how the body compensates to maintain dynamic balance. With the movement of center of mass, reaction forces activate and act on other body joints, particularly those in the lower extremity. Active control of muscles at the hip, knee, and ankle is then required to prevent the body from collapsing. Studying

the gait cycle helps to understand how joint control is achieved (CRUZ-JIMENEZ, 2017).

Considering the widespread chronic conditions and illnesses present in elderly population and the associated social and health costs, it is important for both medical and scientific communities to know more about the “normal” gait pattern of healthy elderly to stablish a valid database that allows comparison with elderly that need special care (PRINCE; CORRIVEAU; HÉBERT; WINTER, 1997). Muscle strength naturally decreases as people age, which affects their ability to walk. These changes are associated to a decrease in balance ability, independent daily activities, and an increased risk of fall. So, it is important to maintain the ability to walk of the elderly (TODA; NAGANO; LUO, 2015).

There are several gait deviations associated with aging: reduced walking speed, decline in medial-lateral hip control, decreased stride length, increased stride width, increase in the stance phase, reduced peak hip extension, increased anterior pelvic tilt, reduced-angle plantar flexion, reduced hip motion, decreased ambulation efficiency, decreased muscle strength, and impaired balance control (CRUZ-JIMENEZ, 2017).

Aging also affects the vertical component of GRF during walking. Vertical GRF provides correlative information about the ability to walk of elderly people. (TODA; NAGANO; LUO, 2015). It has been observed algebraic difference for the vertical component force peaks (midstance, heel contact and toe-off), which suggests that there may be an age-related decrease in vertical oscillations of the center of the gravity, which results in a lower vertical acceleration in the center of gravity for the older adults. It could represent an attempt by older adults to improve economy (energy expenditure) by minimizing vertical displacements in the center of gravity and by reducing the muscular forces that are required to slow down and speed up the body during walking (LARISH; MARTIN; MUNGIOLE, 1988).

Regarding the results from the anterior-posterior GRF data, there is evidence showing a lower peak during push-off (late stance, between pre-swing and toe-off) for elderly when compared with young participants. It suggests that elderly presents a weaker push-off than younger walkers (PRINCE; CORRIVEAU; HÉBERT; WINTER, 1997). There are also findings that it depends on speed, and that both absolute force peaks (deceleration and acceleration phases) are larger for larger speeds. During the deceleration phase, peak force in the older adults was lower than

in the younger adults for both lower and higher speeds. During the acceleration phase there was found an interaction between age and walking speed. There was no difference between the peak forces produced by old and young adults for slower speed. In contrast, there is an age-related decreased in this peak force at the faster speed of walking (LARISH; MARTIN; MUNGIOLE, 1988).

In summary, the peak forces of old and young adults were found to be equivalent at the lower speed and were lower in the older adults at faster speed of walking. It could suggest that the aging motor system attempts to reduce the forces that must be absorbed during the braking phase of stance, even at slow speeds of walking. This dependency on speed of the GRF during the propulsive phase of stance suggests that the aging motor system may be unable to produce levels of force comparable to those of younger adults at fast speeds. This finding reflects an attempt to prevent the musculoskeletal system from experiencing levels of force that are potentially dangerous (LARISH; MARTIN; MUNGIOLE, 1988).

To calculate the moment of force, one should know the segment masses and its moments of inertia. The characteristics of body segments are listed by anthropometric data. It is usually obtained from postmortem studies. A common way to calculate the moment of force is using a so called inverse dynamic approach, which requires the knowledge of the segment inertia and its center of mass position in relation to a bone reference. There are differences between the anthropometric characteristics of young and elderly, and there are tables for the distribution of mass and moment of inertia of body segments (PRINCE; CORRIVEAU; HÉBERT; WINTER, 1997).

In general, the patterns of power profiles are similar between elderly and young participants, however there are some differences when the energy absorbed or generated (power and work) are compared (PRINCE; CORRIVEAU; HÉBERT; WINTER, 1997).

The ankle plantar flexor power during push-off (late stance, pre-swing or toe-off) is higher in the young compared to the fit and healthy elderly. It means that elderly show a less vigorous push-off in the elderly, which have been noticed using both 2D and 3D analysis protocols. Peak ankle power was found as the strongest predictor of step length as it explains more than 52% of the step length variance. The young present a higher work done by plantar flexors when compared with the elderly (PRINCE; CORRIVEAU; HÉBERT; WINTER, 1997). The literature appears to have a consensus supporting a smaller role of the ankle plantar flexors in gait for older

adults, but support for a compensatory action at either the knee and/or hip is weaker (BOYER; JOHNSON; BANKS; JEWELL *et al.*, 2017).

The knee absorption during the transition between stance and swing is higher in the elderly than in young adults. Thereby, the absorption of energy generated by the major push-off phase is greater in the elderly (50%) than in the young (16%). There is a tendency for the peak knee absorption to decrease with respect to age during both mid and late swing phases. It means less demand on the musculature to reduce the angular velocity of the leg and it is directly related to the lower push-off by ankle musculature (PRINCE; CORRIVEAU; HÉBERT; WINTER, 1997).

The hip pull-off power has important contribution (around 16%) to the gait of elderly, however there are divergence on whether this quantity increase or decrease with age advancement (PRINCE; CORRIVEAU; HÉBERT; WINTER, 1997).

Moreover, the literature on elderly people gait showed different distributions of joint moments and powers compared to young individuals. A net extensor pattern of moments at the ankle, knee, and hip joint provides support during the stance phase. Figure 11 shows a representation of age-related changes in lower extremity joint moment. It suggests that the elderly and the young have different support strategies (TODA; NAGANO; LUO, 2015).

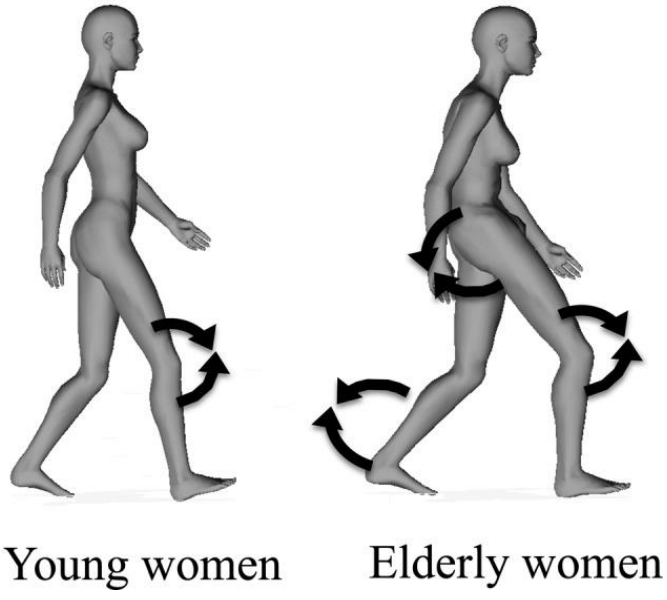



*Figure 11 Joint moments graphical representation of elderly and young females (BOYER; JOHNSON; BANKS; JEWELL et al., 2017)*

The literature suggests that elderly presents lower first and second peak and a higher minimum value at mid-stance (vertical GRF) when compared with young

people, but the results from Toda et al. (2015) does not endorse these assumptions. Only the second peak of the vertical GRF of the elderly women is found to be lower compared with young women (TODA; NAGANO; LUO, 2015).

Healthy young adults asked to walk with a bent-hip, bent-knee posture or crouched gait exhibit a GRF pattern consistent with that presented by older adults: a reduction in the peak vertical and horizontal GRFs. Kinematic results using matched gait speed also provide evidence that older adults walk with a less upright posture or forward leaning posture: more hip flexion at heel strike, greater peak hip flexion, and more knee flexion at heel-strike. It could play a significant role in increasing the metabolic energy cost associated with locomotion in older adults, contributing to a higher energetic demand during locomotion for older adults (BOYER; JOHNSON; BANKS; JEWELL *et al.*, 2017).

Regarding the joint moments, both young and elderly women results show that there is a relation between the knee extension moment and the minimum value of the vertical ground reaction. For the elderly group of females, was noticed a relationship between the hip extension moment and the first peak of the vertical GRF, and the ankle plantar flexor moment was related to the second peak of vertical GRF. (TODA; NAGANO; LUO, 2015).

The most accepted hypothesis is that older adults present lower motion and rely less on ankle. Regarding ankle kinetics, there are indications of a smaller dorsiflexion moment and power generation for older adults (BOYER; JOHNSON; BANKS; JEWELL *et al.*, 2017).

There are mixed conclusions which hinder to stablish a compensation mechanism of hip kinetics (peak joint moment and power generation) with age to explain for the reduction in ankle power generation. The divergence occurs when comparing studies with similar speeds among young and elderly individuals and studies that do not take group speed into account. While results from non-matched gait speeds were not different, the results from studies with matched gait speeds presented greater hip power generation, which could counter the decrease in ankle power generation on the same phase of the gait cycle. So older adults walking at similar speeds to young adults accomplish that by increasing the contribution of the hip (BOYER; JOHNSON; BANKS; JEWELL *et al.*, 2017).

The literature does not support a moderate or large difference in overall knee kinematics in older adults because the size and direction of the effects found are

not consistent among the literature. Only the range of motion of knee present differences. The conclusions differ when comparing matched and unmatched gait speeds. Studies that used matched speeds found more individual gait variables with significant standardized effects. It also found that young adults present greater knee extension at heel-strike and peak flexion in swing (BOYER; JOHNSON; BANKS; JEWELL *et al.*, 2017).

Although there is evidence of reduction in knee extensor function with age, and an association of this reduction with injuries and joint degeneration, no systematic age effects were found for knee kinetics. It leads to the conclusion of the need for more investigation on knee function changes with age, ideally with large well-controlled cohorts (BOYER; JOHNSON; BANKS; JEWELL *et al.*, 2017).

Thus, it suggests a critical role of walking speed on knee kinematics, and that gait speed should be carefully considered when designing a study and when interpreting data to determine the mechanisms of age-related gait differences (BOYER; JOHNSON; BANKS; JEWELL *et al.*, 2017). Moreover, during ambulation, participants use an optimal walking speed, which is a self-selected gait speed at which they are comfortable. At this rate, the body is efficient in its energy consumption (CRUZ-JIMENEZ, 2017). So, for comparison, it is important to match speeds, or be cautious to keep track of the participant speed, so that speeds are similar between participants.

Another aspect of biomechanical studies is that it often has small cohort sizes and not all studies report data in a consistent manner. The fact that discrete time-points are selected does not let us know if there are greater effects at one joint over another or if the changes or if the changes are isolated to specific phases of the gait cycle (BOYER; JOHNSON; BANKS; JEWELL *et al.*, 2017).

# 4. METHODS

## 4.1. Data collection

### 4.1.1. Participants and ethics

Twelve out of fifteen participants that have participated in previous studies at our laboratory, Laboratório de Avaliação e Pesquisa em Atividade Física (LAPAFI), were included in this data exploratory study. They were elderly women (age ≥ 60 years old), without gait pathology and free of lower extremity injuries or surgical history. All participants were informed about the experimental protocol and provided their informed written consent to participate in the study. The approval for data acquisition from the ethical committee of the Pontifical Catholic University of Rio Grande do Sul (PUCRS) in Porto Alegre (Brazil) was received, under CAAE 55674116.3.0000.5336.

We have also worked with an external dataset containing GRF data, which has already been pre-processed and approved by its ethical committee (HORSAK; SLIJEPCEVIC; RABERGER; SCHWAB *et al.*, 2020). The dataset is known as GaitRec, and contains GRF data of healthy controls and gait-impaired participants of both sexes, but we have focused on women healthy controls, to compare with data from our laboratory of older adult women. More details will be described in the following sections.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. The authors declare that there is no conflict of interest.

### 4.1.2. Force Platforms

The force platform used to acquire data for this research is a strain gauge type from BTS Bioengineering manufacturer. The manufacturer provides information about the type of sensors applied in this device as accusphere sensors, as shown in Figure 12, which enables the force platforms to perform measurements across the entire surface of the force platform. In Annex A is attached the force plate formulae provided by the device's company (P-6000, BTS Bioengineering, Italy) with information of force transducers and how the force components are calculated from measured signals. In Annex B, there is a table with specific information of the force platform features provided by the manufacturer website (BTS Bioengineering).


*Figure 12 Force platforms modules and zoom-in of force transductors (accusphere sensors). Source: Modified from BTS Bioengineering user manual and website.*

GRF was measured with a sampling rate set at 1000 Hz. The software provided by BTS Bioengineering allows the user to export GRF data into text files (*.emt). It is a text file containing 11 headlines and 8 columns: Frame, Time, r gr.X, r gr.Y, r gr.Z, l gr.X, l gr.Y, and l gr.Z; r = Right; l = Left; gr = ground reaction force. X = anterior-posterior, Y = vertical, Z = medio lateral. Each time the participant repeats the static or dynamic experiment a new file is generated.

### 4.1.3. Experimental procedure

First, before the walking repetitions, the subject performed a static symmetry assessment (SSA), to be specific, the participant is asked to stand still for five seconds on two force platforms to record the distribution of bodyweight on each foot, as shown in Figure 13. Each volunteer performed walking repetitions in a single assessment session, while they did not undergo any intervention. The number of repetitions varies for each participant, and it will be presented later in the results section. For each trial the GRFs through time were measured, while the subjects walked on a 6m walkway instrumented with four pairs of GRF plates (0.6 x 0.4 m² of sensor area for each force platform – in total there is 2.4 x 0.8 m² of sensitive area to collect GRF data), as shown in Figure 14. The three-dimensional GRF vector components were recorded by eight BTS force platforms (P-6000, BTS Bioengineering, Italy), positioned in pairs close-

together. The laboratory (LAPAFI) environment was kept controlled during the data collection. The subjects were instructed to walk at a self-selected speed.

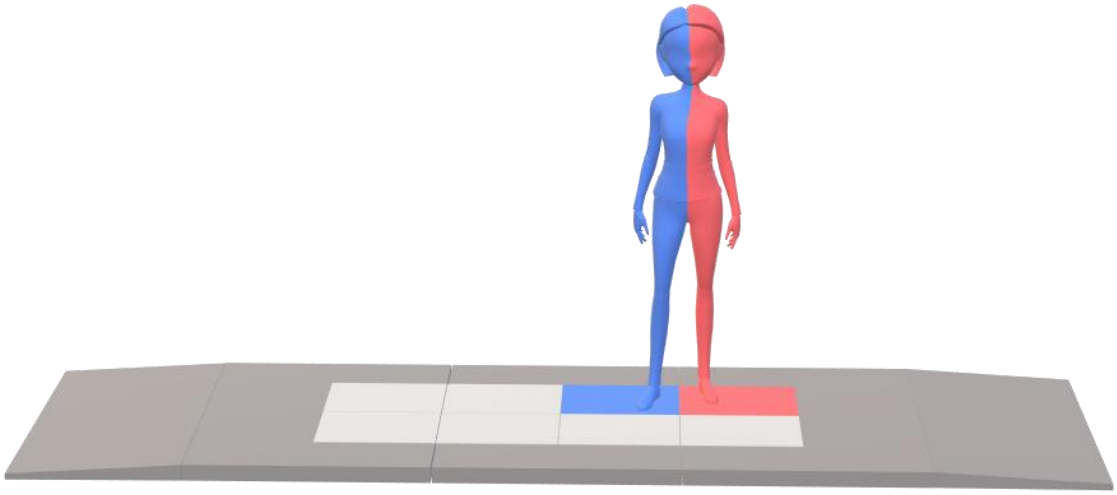


*Figure 13 Graphical representation of static test: stand still for 5 seconds with feet and arms in parallel for GRF recording. Blue and red: right and left, respectively. (The Author, 2021).*

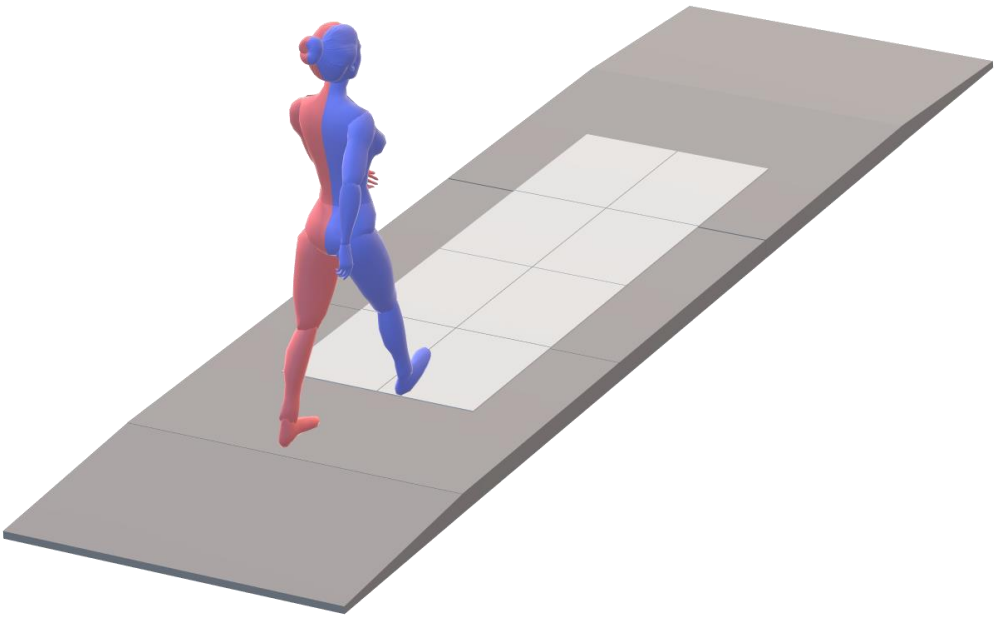


*Figure 14 Graphical representation of dynamic test: walk through with self-selected speed. Blue and red: right and left, respectively. (The Author, 2021).*

### 4.1.4. Experimental datasets: LAPAFI and GaitRec

The LAPAFI dataset for this experiment was built gathering data from the laboratory local data storage. The files are saved in folders for each participant with their names on it. The files are divided into raw data and reports. Raw data are the text files containing bilateral GRF components (three dimensions) for static and dynamic

experiments. The reports are PDF files containing information provided from the manufacturer software that performs the gait analysis.

There was also a third data file, a Microsoft Excel (Office 365) file, containing demographic data which has also been incorporated to the final dataset. It contains age informed, date of birth, height, mass measured, and date of evaluation (data acquisition), which was filled by the former researchers during data collection. In addition to these, we have also calculated the age to compare it with informed age, and we also calculated body weight (from static test). The age was calculated using the date of birth and date of examination, in Microsoft Excel (Office 365) applying the YEARFRAC function. Afterwards, the data are built into a single MATLAB 2020a (NATICK, MASSACHUSETTS: THE MATHWORKS INC.) file (.mat), together with GRF data, to facilitate data processing, resulting in a dataset that is possible to share with other researchers. The names of the participants have been anonymized manually, changing the names to an ID from 1 to 15, preceded by the group identification “elderly” label.

The GaitRec dataset consists of metadata (demographic and experimental condition information, called annotation by the authors), GRFs and center of pressure data for both healthy controls and gait impaired participants (HORSAK; SLIJEPCEVIC; RABERGER; SCHWAB *et al.*, 2020). We have focused on comparing our dataset (LAPAFI) with the preprocessed healthy control (HC) portion of GaitRec dataset. We have only worked on ground reaction data, not on center o pressure at this moment. For this, we have downloaded the .csv files, and selected only healthy control women participants, that walked at self-selected speed and with usual shoes, aiming to meet similar experimental conditions.

## 4.2. **Data Preparation**

### 4.2.1. GaitRec GRF data

GaitRec is a large-scale GRF dataset of healthy and impaired gait. It provides gait measurements and phenotypic annotation. The technology type is based on force sensor and visual observation methods. The factor types presented are experimental condition, musculoskeletal impairment, age, sex, shod condition, and walking speed. The dataset contains gait measurements from humans in a laboratory environment, specific details about their experimental protocol are presented on Annex C. The

GaitRec dataset comprises completely anonymized GRF measurements from 2085 patients with different musculoskeletal impairments (“gait disorders”) and data from 211 healthy controls including additional metadata such as age, sex, shod condition, walking speed condition, and etc. (HORSAK; SLIJEPCEVIC; RABERGER; SCHWAB *et al.*, 2020).

We have worked on the ready to use processed files. From the 211 healthy controls participants we have selected only females, resulting in 105 female participants, from this we have selected the GRF curves collected with usual shoes, at self-selected speed. The number of curves vary between participants, so we have calculated the mean curves of each foot and the mean of both feet together, which are used later for 0D and 1D analysis. We divided the participants into three groups based on age: young (15 to 29 years old), adult (30 to 49 years old) and older adult (greater than 50 years old). From these groups we have selected twelve women of each group to perform the discrete and continuous data analysis. This selection was done randomly using MATLAB 2020a functionalities that allows for reproducibility: data sample and RandStream functions, which is available with the MATLAB 2020a code provided in this study on Appendix A. The reduction to twelve participants had to be done to allow us to run the statistical analysis, but unpaired data can also be compared, as shown by the software examples.

### 4.2.2. LAPAFI GRF data

The data preparation (pre-processing) is important before conducting analysis on data. It is done to mitigate errors in the final analysis to ensure the quality of the data being analyzed. The approach of pre-processing depends on the type of the data, on how it has been organized previously, and on the type of analysis that will be performed.

Prior to conducting either the discrete (0D) or the continuous analysis (1D), we have performed a series of processes to prepare the data. These processes are presented in a simplified flow diagram in Figure 15. Each function is explained in detail in the following sections. Appendix A presents the complete flow diagram including the functions, outputs and inputs together with the connections between them.

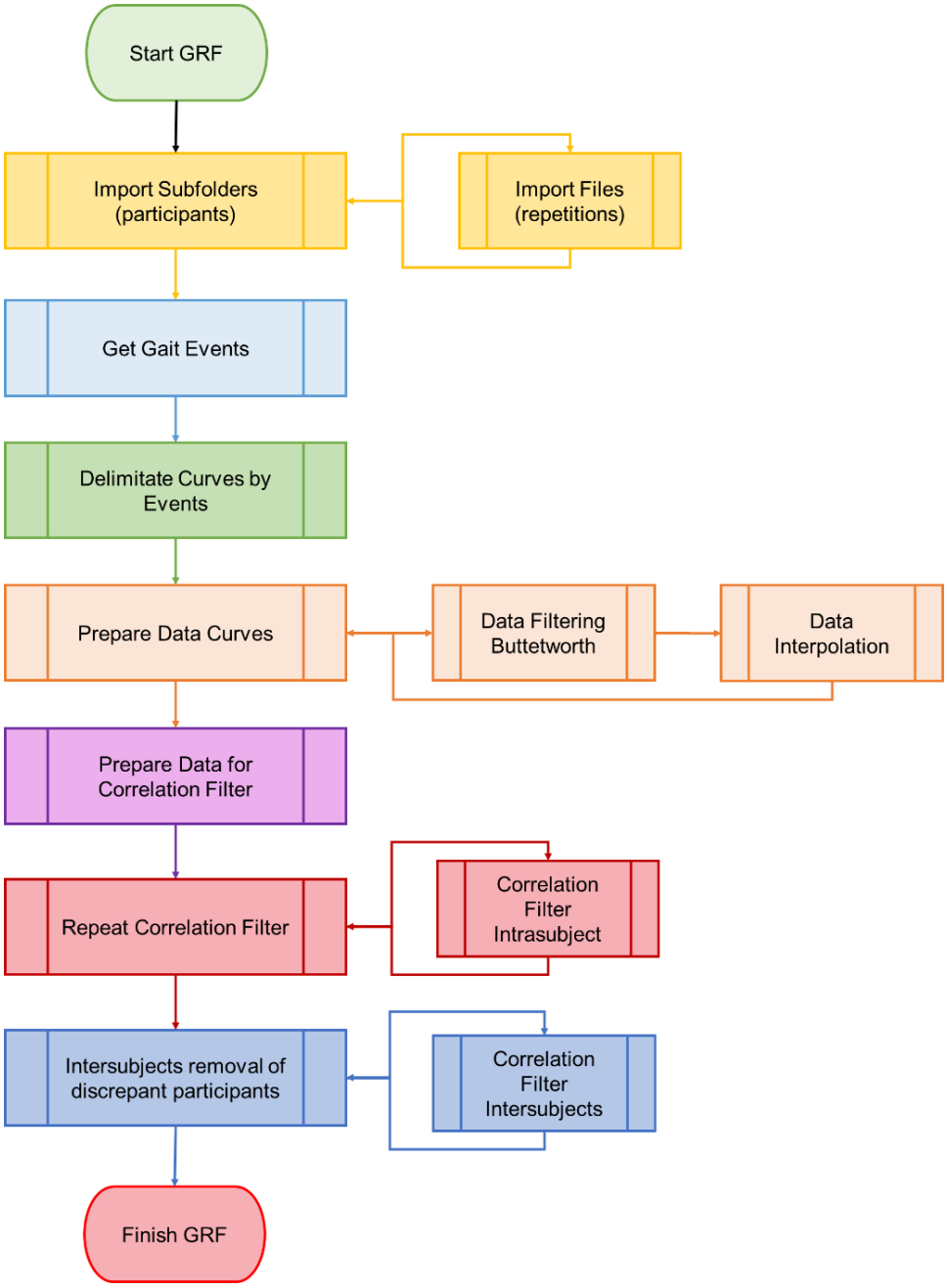



*Figure 15 Simplified main script flow diagram for GRF data preparation (The Author, 2021).*

a) Importing data into MATLAB environment

First, before importing GRF data into MATLAB 2020a environment, we have loaded a previously prepared data structure called elderly_metadata.mat, which contains the following information: ID, DOB, DOE, AGE_calculated, AGE_informed, MASS_kg, and HEIGHT_m.

We have created two functions to load the GRF data from our participants into a single data structure. The function importSubfolder.m calls the function importFile.m to mount the dataset into a single data structure called Data_GRF that contains three

columns for each participant. The first column contains the data from the static test in a table with seven columns: time stamp, GRF_AP_R, GRF_V_R, GRF_ML_R, GRF_AP_L, GRF_V_L, and GRF_ML_L, which is used later to calculate the body weight and body mass of the participant. The second column contains all dynamic test files concatenated into a single table, containing the same seven columns described above. The third column contains the number of dynamic test files loaded for that participant (this is not an equal number and refers to the number of times the participant walked through the walkway). The importSubfolder also returns the sampling frequency (Fs) which is later used as input in the prepDataCurves.m function.

We have then changed the signals of the following variables: GRF_AP_R, GRF_AP_L, and GRF_ML_L. So, both Anterior and Medial forces are always shown as positive values. It is necessary because the walking sense is opposite to the system's origin coordinates and because right and left foot shows opposite senses relative to each other. Then we plotted the concatenated GRF_V for both right and left foot, for general visualization.

b) Body weight calculation from static data

Before moving to the event detection, we have calculated the body weight (BW) (N) and mass (kg) of each participant. Following the parallel system of forces, we have collected static data during standing for around five seconds for each participant, then we calculated the mean of the summed vertical forces on each foot, resulting in the body weight:

$$BW = \frac{\sum_{i=1}^{n} GRF\_V\_R_i + GRF\_V\_L_i}{n}.$$

where $GRF\_V\_R_i$ and $GRF\_V\_L_i$ are the values of GRF of vertical component of right $R$ and left $L$ foot at the time stamp $i$, and $n$ is the total number of time stamps recorded. To calculate the body mass, we have divided the body weight by the value of the gravitational acceleration on earth 9.807. Both body weight and body mass information are incorporated into the previously loaded data structure called elderly_metadata, to be used later in the prepDataCurves.m function.

c) Detect gait cycle events and delimitate curves

The event detection collects where each stance is supposed to start and end in the time domain. We have created a function called getEvents.m which reads the vertical component of GRF and detects if it is a number or not. Then we created a set of conditional statements to detect the begin and the end of the stance. The begin as

the transition of the instant where the previous data point was not a number. To detect the end of the stance as the transition of the instant on which the next data point was not a number, the function registers these indices (events) as well as the number of stances detected (n_steps). The delimitation of the curves is done using a function called delimitateCurves.m. It simply retrieves the data for each stance phase, foot and GRF component receiving as parameters the original data structure, the number of stance phases detected, the events (the begin and the end of each stance phase), and the GRF components' indices to work on.

d) Preparing data curves: body weight, noise removal filtering, downsampling, interpolation and primary cleaning curves

The function called prepDataCurves.m is the first filter applied on the GRF data. It receives as parameters the delimited curves, the number of stances detected, the weight, the sampling frequency and the cut-off frequency. This function was designed to normalize by body weight, remove too short curves, smooth data using a digital filter, down sample from 1000 Hz to 100 Hz, and interpolate the curves to 101 points.

Normalization techniques are used to compare gait data between individuals of significantly different heights and/or bodyweights; GRF data is often normalized to body weight (CHOCKALINGAM; HEALY; NEEDHAM, 2018). So, first the function divides all components of the GRF vector by the body weight calculated for each participant using the static tests data. Then we set an empirically defined exclusion criterion that if the length of the stance is smaller than the sampling frequency divided by five (200 milliseconds) the data must be discarded. It was necessary for the following steps, as the interpolation after down sampling would not accept too small vectors.

Noise is the deviation of the sampled data points from the true signal. There are many sources of noise in biomechanics. The noise is also amplified during numerical differentiation. There are efforts to minimize these effects prior to digitization of the signal. Despite the precautions during data acquisition the sampled data will still have some level of noise. A portion of this noise may be removed using smoothing techniques. The type of the filter can vary, according to the desire of removing a certain range of frequencies. We choose a low pass type of filter, which attenuates high frequency components and leave low frequency components unaffected (EDWARDS; DERRICK; HAMILL, 2018).

There are several ways to choose the appropriate cut off frequency for a particular smoothing routine, and it can be a challenging task. For most biomechanical measurements, there is a considerable overlap between the true signal and noise in the frequency domain. Which makes the task of choosing the appropriate cut off frequency an optimization problem with the objective of removing the noise while preserving as much of the signal as possible (EDWARDS; DERRICK; HAMILL, 2018).

The literature shows a wide range in the choice of an appropriate cut-off frequency. There is a recommendation to avoid cut-off frequencies below 20 Hz, as 20 Hz seems to be a good trade-off between reducing noise and preserving as much physiological frequency as possible (HORSAK; SLIJEPCEVIC; RABERGER; SCHWAB *et al.*, 2020).

We have applied a fourth order Butterworth filter, as a dual-pass second order Butterworth filter, as recommended by the literature. By passing the data through the filter twice, the order of the filter is effectively doubled, and the phase lag of the filter becomes zero. We have applied a low-pass fourth order Butterworth filter with zero lag (we passed a second order filter twice), with a cut-off frequency of 20 Hz.

Down sampling is the procedure of reducing the size of data. We down sampled from 1000 Hz to 100 Hz using a built in function from MATLAB 2020a called down sample which receives the data and the factor of reduction, in this case: $\frac{F_s}{100}$, which results in a factor of reduction equal to 10, as our $F_s = 1000$.

Interpolation is the process of estimating an unknown value between two known values. We interpolated the down sampled and filtered curves to 101 points using the self-written function datInterp.m which receives the data, the number of resulting points and the method of interpolation. We have chosen the ‘pchip’ (Piecewise Cubic Hermite Interpolating Polynomial (PCHIP) method which is a shape-preserving piecewise cubic interpolation.

e) Correlation Filter

Clean foot strike: one of the potential problems that may appear while measuring GRF are the incomplete or dirty foot-strike. It means that the foot must touch inside the force platform boundaries, as shown in Figure 16, otherwise not all the GRF will be recorded. Another common problem is when during double stance both feet touch the same platform resulting in a meaningless mixture of the GRF of both sides (KIRTLEY, 2006).

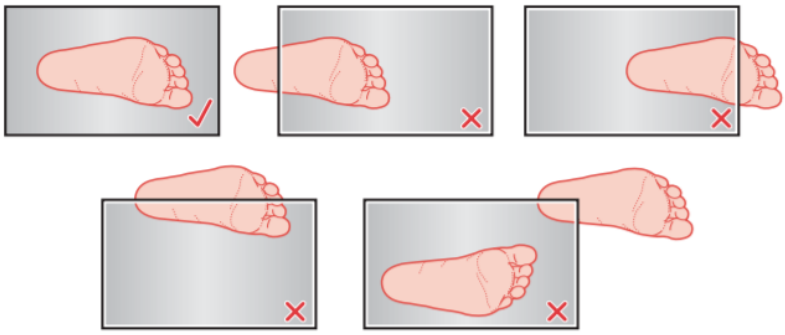


*Figure 16 Representation of possible foot landing position on force platform (KIRTLEY, 2006).*

To overcome this issue of clean strike one option is to repeat the experiment until there are only clean strikes recorded, but this should be defined before data acquisition. To overcome this issue from a perspective of retrospective data analysis we have developed a filter based on linear correlation between curves. This filter keeps only those curves that present a mean correlation coefficient that surpasses a empirically set correlation threshold. This process is implemented through two functions: repeatCorrFilter.m and corrFilterIntra.m

Before applying the correlation filter we have to prepare the data, so the functions can work on it, we have created a function that does it and it is known as prepDataCorr.m. It simply receives the data that has been prepared (normalized, filtered, downsampled and interpolated) and rearranges it, so calculations like the correlation coefficient can be done.

The repeatCorrFilter.m function receives the three components of GRF data in the format of a cell containing all curves of each participant and the limiar correlation (corr_limiar) the user wants the data to achieve. It returns as output the optimized data for all participants (output) and the total number of repetitions (rep_n) of the correlation filter to achieve the desired correlation threshold.

This function calls other function called corrFilterIntra.m, and it has the purpose of making the filtering process recursive. It means that the output of the first round of filtering is the input of the next round of filtering until there is no more necessity of filtering. It means: while the repeat status returned by corrFilterIntra.m function returns true, the number of repetitions is incremented and the repetition status is refreshed until all repeat status of all participants become false and the while loop is finished, as shown in Figure 17.

Finally, the output matrix is mounted using the resultant optimized matrix for the latest repetition incremented (rep_n) when the while loop was finished for each participant. It also returns the dimensions of the resultant optimized matrices (the number of curves resultant).

The corrFilterIntra.m function aims to filter the intrasubject curves. It receives a matrix for each participant containing the number of curves and the 101 points of the three GRF vector components, but the mean correlation coefficient is calculated using only the vertical component. The function also receives the correlation threshold that is the mean correlation coefficient the user wants all participants' curves to achieve. It is defined empirically in conjunction with visual inspection of the survived curves.

The output of this function is an optimized matrix containing all curves that have achieved the correlation coefficient threshold or containing the curves minus the smallest mean correlation coefficient curve. It depends on the repeat status that is also an output of this function. The repeat status (true or false) is the condition for the decision of the corrFilterIntra function to be repeated or not on the repeatCorrFilter function, as shown in Figure 17.

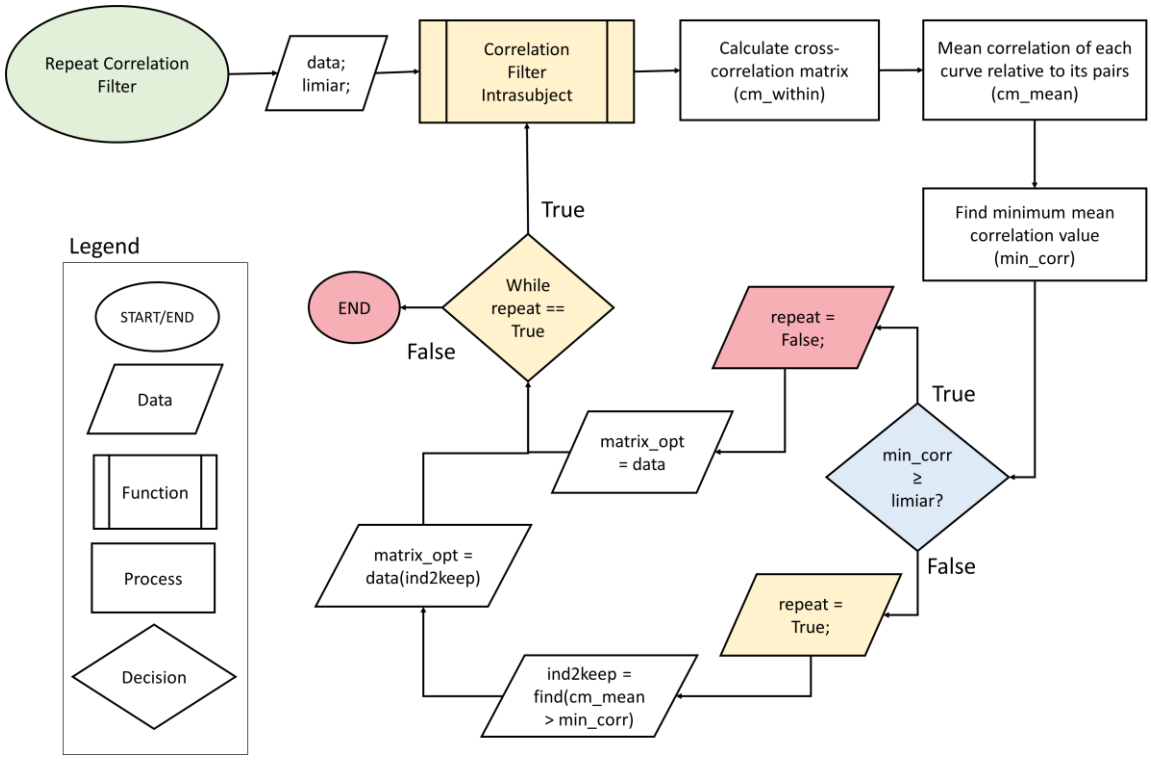



*Figure 17 Flow diagram of the correlation filter process*

Inside the intrasubject correlation filter, we start by calculating the cross-correlation matrix of all curves of a participant, this correlation matrix is calculated using the most common linear correlation coefficient, the Pearson's linear correlation coefficient. Values of the correlation coefficient range from -1 to 1. A value of -1 indicates perfect negative correlation, while a value of +1 indicates perfect positive correlation. A value of 0 (zero) indicates no correlation between the columns. So, the smaller the correlation coefficient's module, the weakest the relationship between the curves.

Then we calculate the mean correlation of each curve compared with its pairs, and finally we find the minimum value, which corresponds to the curve that has, overall, smaller correlation than their neighbors. This value is compared with the correlation coefficient threshold (limiar) set by the user (empirically and supported by visual inspection). If the minimum value of mean cross-correlation is greater than the threshold set by the user the function returns false for the repeat status, and returns the same curves entered as input.

Otherwise, if the minimum value of the mean cross-correlation is smaller than the threshold set by the user, the function returns true for the repeat status, and keeps only those curves that present a mean cross-correlation greater than the minimum mean cross-correlation value. This matrix is updated every time it passes through the corrFilterIntra.m function, until the repeat status becomes false for all participants.

f) Data Exporting and Intersubjects Filtering

Export data: we have mounted into the initial metadata_GRF data structure the data that has been processed in this work. For both right and left foot, and all three GRF vector components we have exported the following data: GRF concatenated (all trials data), GRF delimited (prepared curves), and the resultant curves of intracorrelation filter process using the different correlation coefficient thresholds (85%, 90%, and 95%). Plot number of curves for corrFilter results. To visualize how the correlation filtering process affected the number of registered curves we have plotted the number of curves for each situation. Then we removed participants with less than four curves, to be able to proceed with SPM analysis we have kept only those participants that remained with more than four curves after the 95% correlation filter intraparticipant.

Inter subjects filtering: this step was designed to deal with the interparticipants curves discrepancy. The function interSubjFilter.m is applied to handle this issue. First it calculates the mean curve of each participant for each foot, then we have called a modified version of the correlation filter, which is called corrFilterInter.m. It had to be modified, because now we are going to use the mean curve of each participant to compare the curves between them, instead of doing it for multiple curves for each participant as in the intra participant case.

We have called the function twice, first with a correlation threshold of 85% and the second round with a correlation threshold of 90%. It has been defined empirically supported by visual inspection, so we do not lose too many participants. Then we mounted a new data structure called elderly_subj_final containing only those participants that survived the interparticipants correlation filtering processes. We have also calculated the mean curves between right and left foot, that will also be used in the discrete and continuous data analysis. Finally, we saved the elderly_subj_final data structure into an .mat file.

## 4.3. **Traditional Discrete GRF Analysis**

### 4.3.1. Peaks and Impulses Extraction

The peaks and impulses of each participant for both the LAPAFI and the GaitRec datasets were extracted using a self-written function called peaks_impulsesFun.m. The function extracts the discrete parameters from the mean GRF curves between right and left foot and stores it on a data structure that is then exported (saved) for analysis. These are presented on Appendix C. These discrete values were extracted following Vaverka's system of gait analysis based on GRF assessment, but due to limitations on accurately defining the peaks on ML GRF component only one point of each GRF component was chosen, as shown in the following section (F. VAVERKA, 2015).

### 4.3.2. Discrete analysis: Independent Hotelling $T^2$ test and post-hoc t-test

Using the code provided by SPM1D package, we wrote two MATLAB 2020a functions to do the independent Hotteling's $T^2$ test for peaks (force values) and impulses. The independent Hotelling's $T^2$ test is analog of the independent t test. For the multivariate Hotelling's $T^2$ tests, there are $p>1$ variables and their correlations are

important (OLIVE; CHERNYK, 2017). From those discrete values extracted from GRF data presented on Appendix C, we have chosen three scalars from the two datasets (LAPAFI and GaitRec) to compare between groups, as shown in Table 4.

*Table 4 Discrete values chosen for traditional scalar analysis*

| **Force** | **Impulse** |
|---|---|
| min_AP_RL; | total_imp_AP_RL |
| peak1_V_RL; | total_imp_V_RL |
| valley1_ML_RL; | total_imp_ML_RL |

The calculation of independent Hotelling's $T^2$ is presented in the following section for the spm1d. After the Hotelling $T^2$ test, we conducted post-hoc test using independent t tests. We conducted one test for each scalar, we performed N=3 tests on each dataset, between each group being compared, where N is the number of extracted scalars. To retain a family-wise Type I (false positive) error rate of $\alpha = 0.05$ we adopted Šidák thresholds of $p = 0.0170$, where the Šidák threshold is calculated as

$$p_{critical} = 1 - (1 - \alpha)^{\frac{1}{N}} = 1 - (1 - 0.05)^{\frac{1}{3}} = 0.0170.$$

The groups were compared according to the age classification, as shown in Table 5.

*Table 5 Groups compared using independent Hotellings $T^2$ for three scalars and post-hoc t-test with Šidák corrected p-critical (p=0.0170)*

| Data A | Data B |
|---|---|
| GaitRec.young | GaitRec.adult |
| GaitRec.young | GaitRec.olderAdult |
| GaitRec.adult | GaitRec.olderAdult |
| LAPAFI | GaitRec.young |
| LAPAFI | GaitRec.adult |
| LAPAFI | GaitRec.olderAdult |

We aim to compare the results from traditional scalar extraction analysis with the equivalent vector field analyses, which is presented in the following section.

## 4.4. SPM Application on Experimental GRF

We have created a function to run the hypothesis tests based on the software package examples, which were previously developed and is publicly available at www.SPM1D.org. The function is named hotellingT²_posthoc.m. The MATLAB codes and data are available following the instructions on Appendix B. The groups compared are presented on Figure 18.

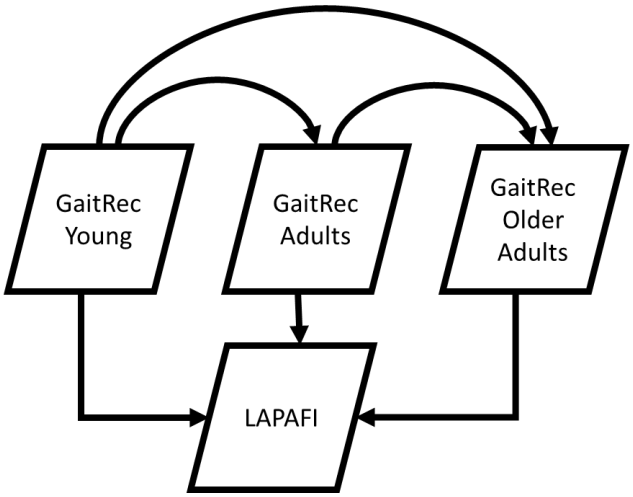



*Figure 18 Groups compared using independent Hotelling's T² on vector-fields of GRFs.*

On the next section we present the results from data preparation, scalar extraction, discrete analysis (hypothesis testing on scalars points), and SPM1D (hypothesis testing on vector-fields).

# 5. RESULTS

## 5.1. Participants

### 5.1.1. LAPAFI

The result of data preparation yielded twelve participants, which the mean curves, survived the corrFilterInter.m filtering function, their characteristics are presented on Table 6.

*Table 6 Characterization of LAPAFI participants*

| ID | Age (years) calculated | Age (years) informed | Mass (kg) measured | Mass (kg) calculated | Height (cm) | Weight (N) |
|---|---|---|---|---|---|---|
| 1 | 67 | 71 | 68 | 67 | 147 | 658 |
| 3 | 60 | 60 | 85 | 82 | 151 | 801 |
| 4 | 68 | 79 | 63 | 64 | 148 | 625 |
| 5 | 65 | 66 | 93 | 93 | 159 | 913 |
| 6 | 63 | NaN | NaN | 61 | NaN | 595 |
| 7 | 66 | 66 | 61 | 60 | 148 | 594 |
| 9 | 63. | NaN | NaN | 84 | NaN | 824 |
| 11 | 66 | 66 | 52 | 53 | 151 | 525 |
| 12 | 82 | 82 | 79 | 79 | 157 | 779 |
| 13 | 82 | 82 | 58 | 58 | 146 | 567 |
| 14 | 69 | 69 | 59 | 58 | 151 | 565 |
| 15 | 65 | 65 | 51 | 59 | 158 | 577 |
| **Mean** | 68 | 71 | 67 | 68 | 152 | 669 |
| **SD** | 7 | 8 | 14 | 13 | 4 | 126 |

SD: standard deviation. Both Mean and SD were calculated without NaN (not a number) values.

### 5.1.2. GaitRec

The random selection of twelve participants for each age group: young (15 to 29 years old), adult (30 to 49 years old), and older adult (greater than 50 years old) from the GaitRec dataset resulted in three groups with the following characteristics presented in Table 7, Table 8, Table 9.

*Table 7 Characterization of GaitRec young participants*

| ID | Age (years) | Weight (N) | Mass (kg) | Height (cm) |
|---|---|---|---|---|
| 204 | 16 | 563 | 57 | 163 |
| 115 | 22 | 633 | 64 | 175 |
| 151 | 23 | 603 | 61 | 171 |
| 113 | 23 | 654 | 67 | 176 |
| 196 | 25 | 659 | 67 | 162 |
| 149 | 26 | 717 | 73 | 165 |
| 184 | 26 | 616 | 63 | 154 |
| 136 | 26 | 537 | 55 | 165 |
| 49 | 27 | 677 | 69 | 175 |
| 101 | 27 | 761 | 78 | 174 |
| 4 | 28 | 553 | 56 | 175 |
| 81 | 29 | 567 | 58 | 162 |
| **Mean** | 25 | 628 | 64 | 168 |
| **SD** | 4 | 69 | 7 | 7 |

SD: standard deviation

Although the young group has participants from 15 to 29 years old the random selection did not sort participants with 15 years old.

*Table 8 Characterization of GaitRec adult participants*

| ID | Age (years) | Weight (N) | Mass (kg) | Height (cm) |
|---|---|---|---|---|
| 129 | 30 | 501 | 51 | 163 |
| 139 | 32 | 556 | 57 | 172 |
| 143 | 37 | 669 | 68 | 178 |
| 26 | 37 | 633 | 64 | 163 |
| 112 | 38 | 577 | 59 | 170 |
| 188 | 40 | 883 | 90 | 172 |
| 181 | 40 | 567 | 58 | 152 |
| 24 | 44 | 668 | 68 | 173 |
| 15 | 44 | 535 | 54 | 160 |
| 116 | 46 | 699 | 71 | 160 |
| 85 | 47 | 780 | 80 | 165 |
| 189 | 49 | 617 | 63 | 172 |
| **Mean** | 40 | 640 | 65 | 167 |
| **SD** | 6 | 109 | 11 | 7 |

SD: standard deviation

*Table 9 Characterization of GaitRec older adults' participants*

| ID | Age (years) | Weight (N) | Mass (kg) | Height (m) |
|---|---|---|---|---|
| 17 | 50 | 617.1 | 62.9 | 1.69 |
| 211 | 51 | 629.7 | 64.2 | 1.55 |
| 6 | 51 | 786.4 | 80.2 | 1.71 |
| 30 | 51 | 627.5 | 64.0 | 1.64 |
| 95 | 52 | 1060.5 | 108.1 | 1.70 |
| 16 | 52 | 990.0 | 100.9 | 1.66 |
| 193 | 54 | 542.4 | 55.3 | 1.59 |
| 187 | 55 | 929.5 | 94.8 | 1.73 |
| 172 | 55 | 658.9 | 67.2 | 1.67 |
| 98 | 55 | 536.5 | 54.7 | 1.61 |
| 68 | 55 | 599.9 | 61.2 | 1.72 |
| 42 | 67 | 770.2 | 78.5 | 1.60 |
| **Mean** | 54 | 729.1 | 74.3 | 1.66 |
| **SD** | 4.5 | 178.2 | 18.2 | 0.06 |

SD: standard deviation

In the following sections we present the data preparation results (LAPAFI dataset) and data analysis (LAPAFI and GaitRec).

## 5.2. **Data Preparation LAPAFI**

### 5.2.1. Concatenated Data

The results of concatenating all dynamic trials for each participant are presented on Figure 19 and Figure 20, as well as the number of trials concatenated for each of the initial 15 participants. It is before delimitating and performing any process of filtering on the curves, so it is possible to visualize discrepant curves. Here we present only the vertical GRF component curves concatenated for general visual inspection.

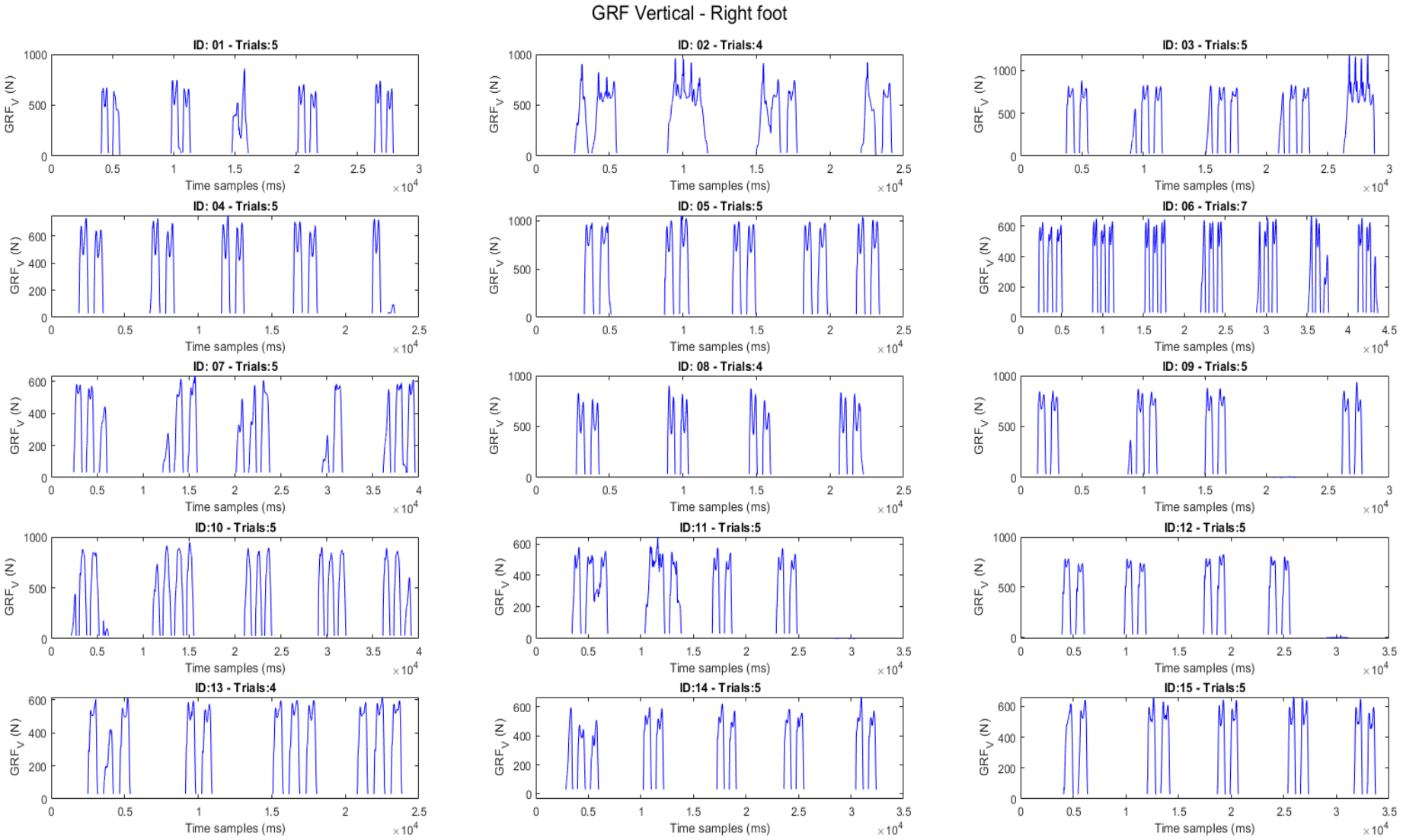



*Figure 19 Concatenated trials right foot vertical GRF component (The Author, 2021).*

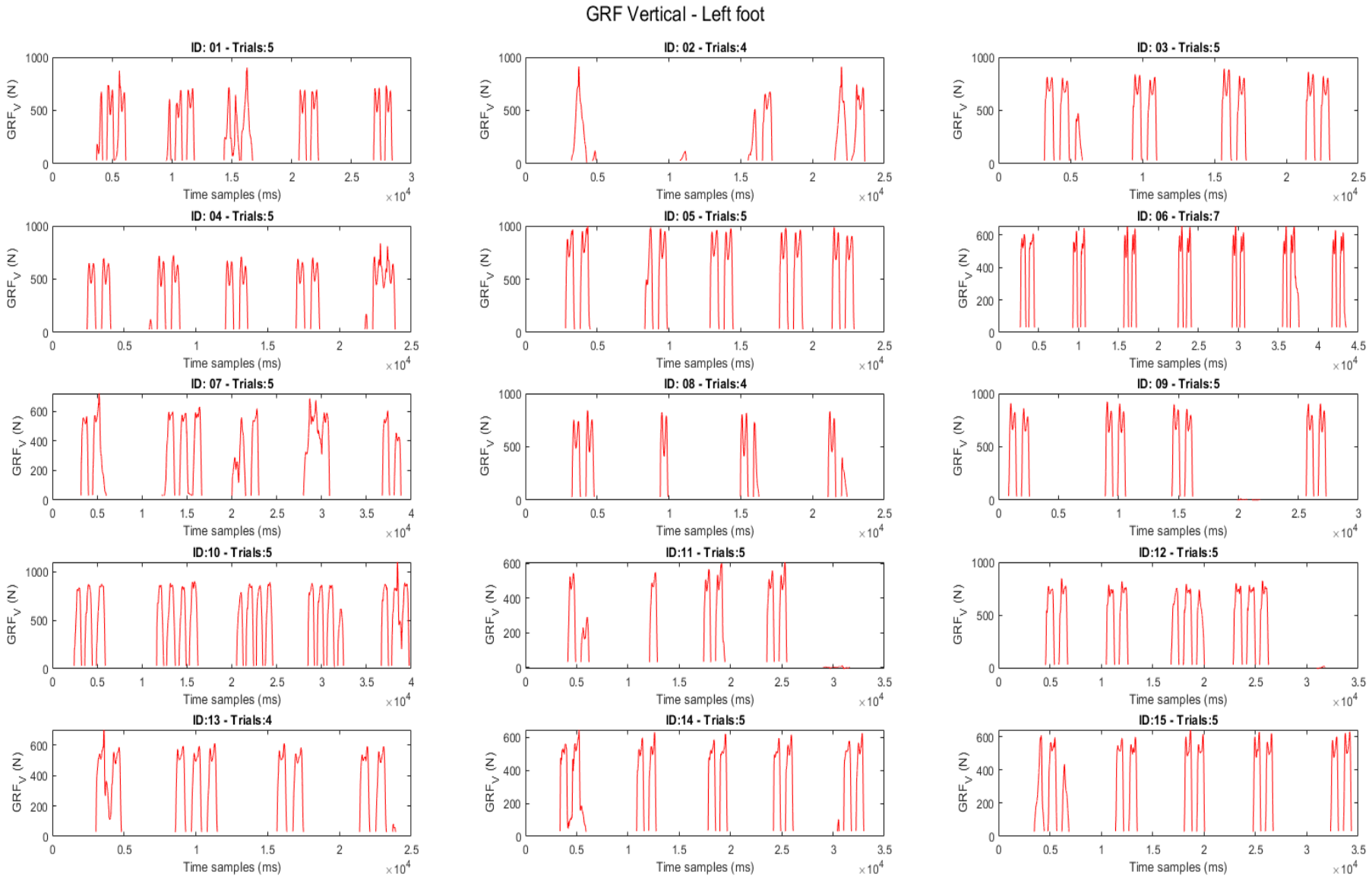



*Figure 20 Concatenated trials left foot vertical GRF component (The Author, 2021).*

### 5.2.2. Delimited Curves Before Correlation Filter

The prepared curves for correlation filtering are presented in Figure 21 and Figure 22, for the vertical GRF vector component. The shear GRF components (anterior-posterior and medial lateral) are presented on Appendix D. These curves have been delimited, filtered for noise reduction (Butterworth low-pass Filter), down sampled, and interpolated, as specified in methods section. The number of curves registered for each participant (ID) is different, as well as their shapes.

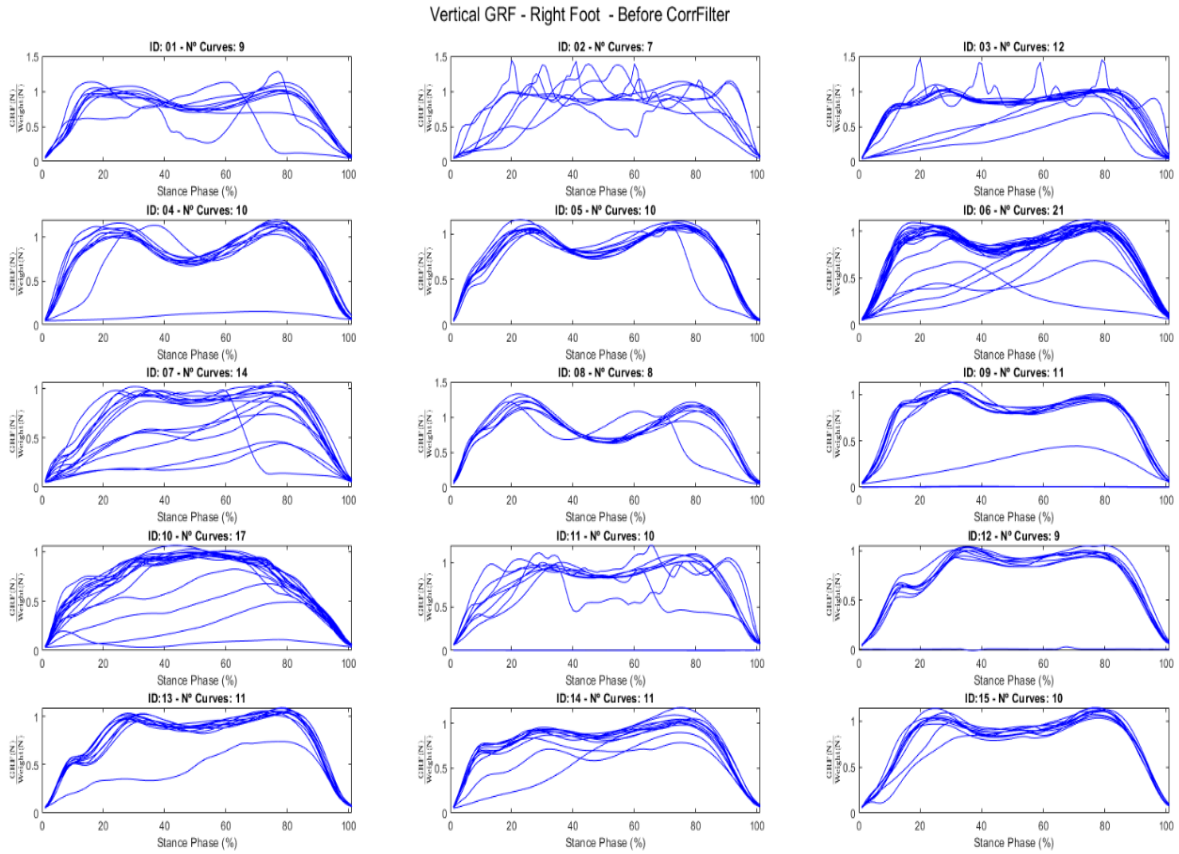



*Figure 21 Vertical GRF component of right foot. Prepared data – before correlation filter (The Author, 2021).*

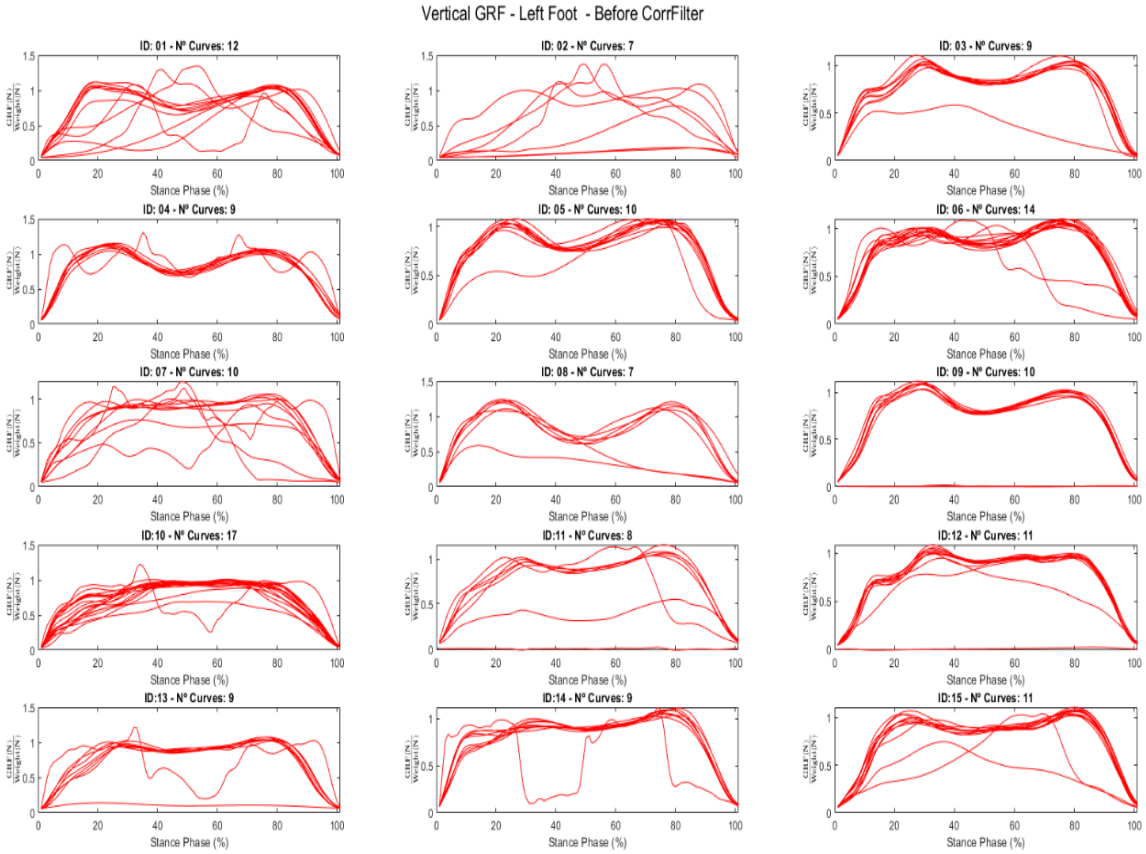



*Figure 22 Vertical GRF component of left foot. Prepared data – before correlation filter. (The Author, 2021).*

There is a non-negligible number of curves that are visually discrepant from their pairs, when comparing it in an intrasubject perspective. To solve this, we developed the correlation filter, explained in more details in the method section.

### 5.2.3. Correlation Filter Intrasubject

The results of the correlation filter intrasubject with a correlation coefficient threshold set as 95% are presented on Figure 23 and Figure 24, while the shar GRF components are presented on Appendix E. It does not mean that there is 95% of correlation between the curves, but that the mean linear correlation of each curve (with the remaing curves of the same participant) that survived the filter has achieved at least 95% of linear correlation. Visually, there is an improvement on discrepancy between the survival curves of each subject.

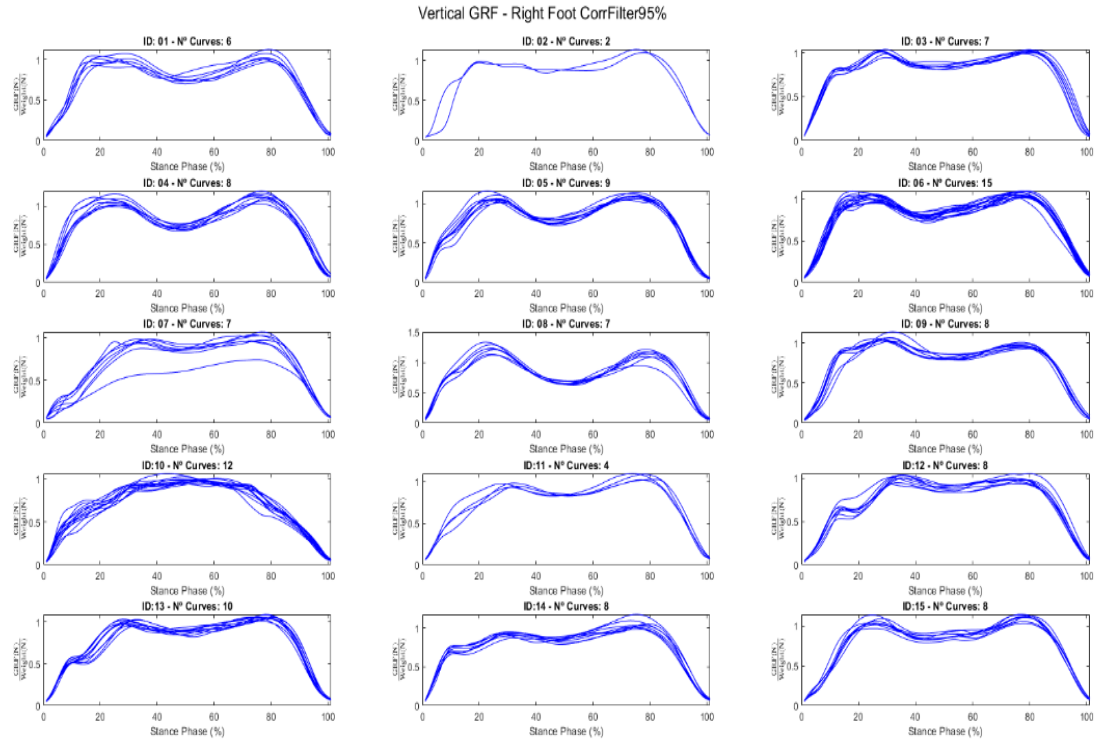



*Figure 23 Vertical GRF after correlation filter (threshold of 95%) right foot (The Author, 2021).*

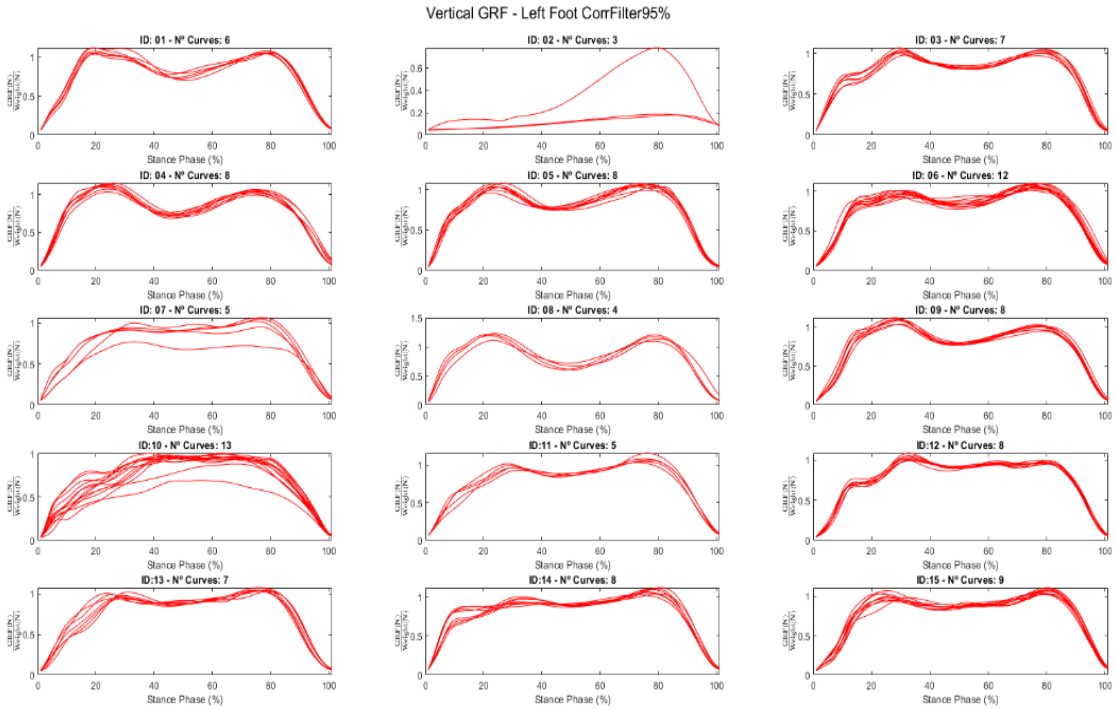



*Figure 24 Vertical GRF after correlation filter (threshold of 95%) left foot (The Author, 2021).*

The participant (ID=2) showed a small number of curves and inconclusive curves, so we had to drop the data to ensure a reliable dataset.

### 5.2.4. Mean curves and Intersubjects Correlation Filter

After removing participants with less than four curves (participant ID = 2, to be more specific), we have plotted the mean curves and from the fourteen mean curves we have removed those participants with curves that were too discrepant between participants, resulting in twelve participants. This process was explained in the method section, for shortness, we present the final mean curves of the twelve participants for Vertical GRF vector component and each foot on the following figures (Figure 25 and Figure 26). The remaining GRF shear components are presented on Appendix F.

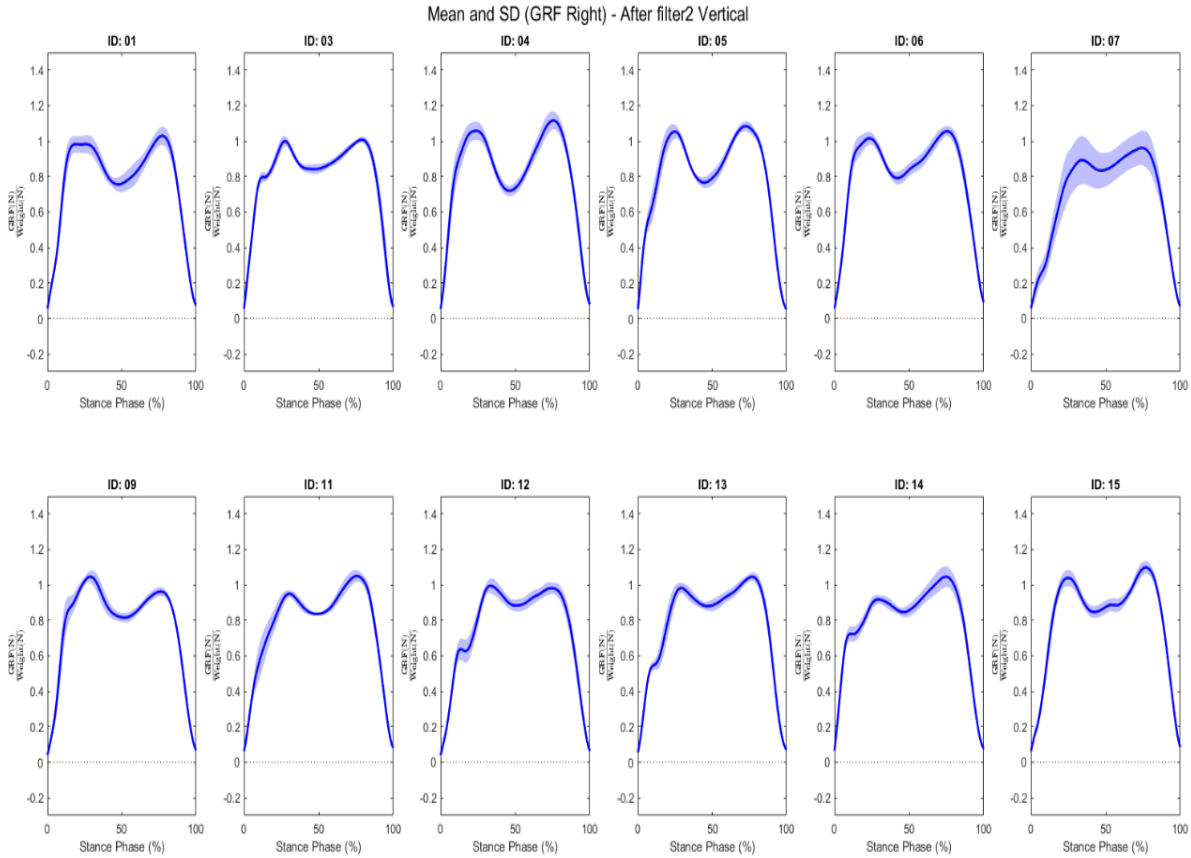



*Figure 25 Vertical GRF mean curves after intersubjects correlation filter, right foot (The Author, 2021).*

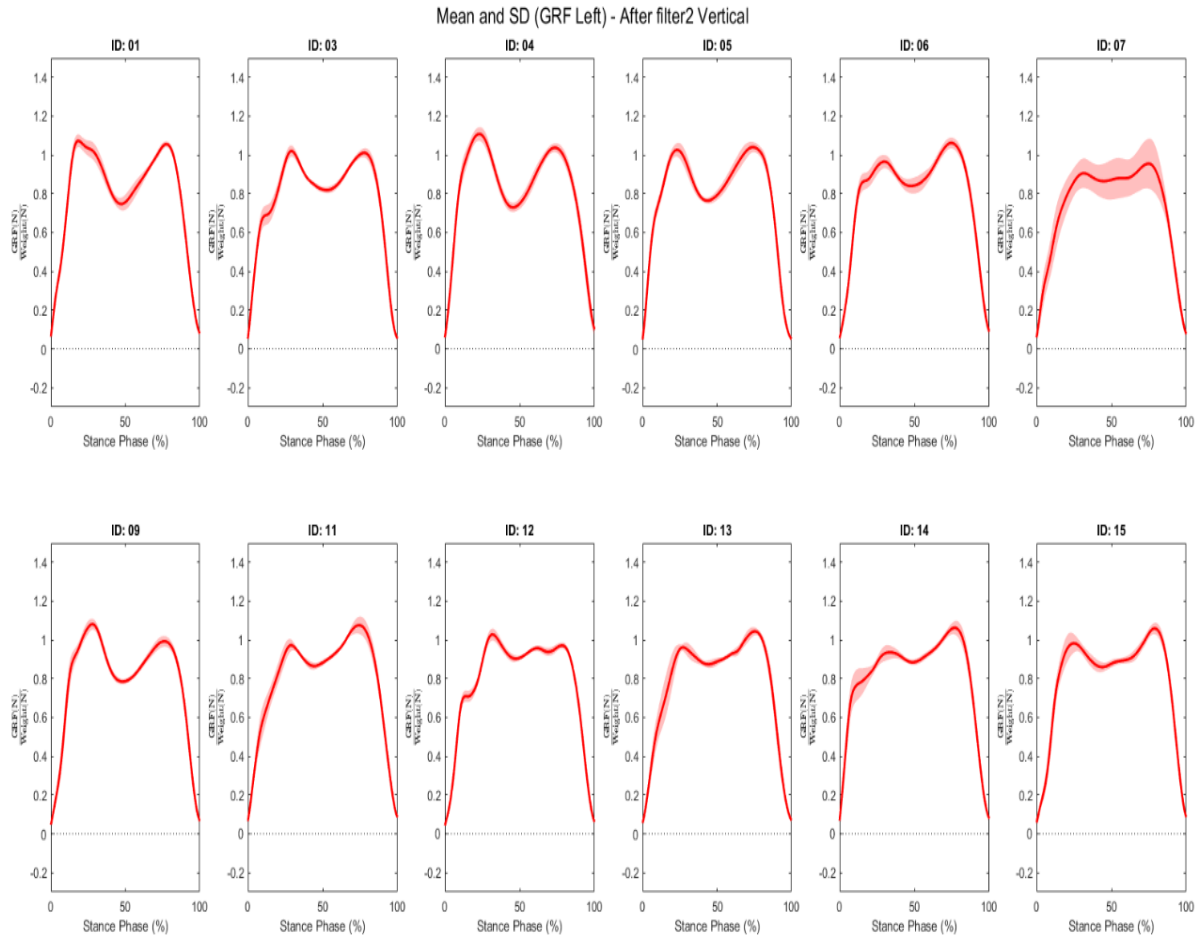



*Figure 26 Vertical GRF mean curves after intersubjects correlation filter, left foot (The Author, 2021).*

In the following section we present the results of traditional scalar extraction and traditional statistical analysis.

## 5.3. **Traditional Discrete Analysis**

### 5.3.1. Peak and valleys

#### a) LAPAFI

The peaks and impulses were extracted from the mean curves between right and left foot of each participant. All curves are presented as the mean curve between right and left foot. For visual inspection, the peaks are presented on Appendix G as red asterisks for each GRF component. In the following section we present the numerical results for the first peak (or valley depending on the GRF component). We also present the results of impulse (numerical integration). We had to choose the first

peak or valley, because medial-lateral component presents such variation that makes it hard to find other peaks precisely.

b) GaitRec

The peaks for visual inspection as red asterisks of the groups: young, adults and older adults from the GaitRec dataset for all GRF components and the twelve randomly selected participants of each group are presented on Appendix H. In the following section we present the numerical values for the first peak (or valley depending on the GRF component). We also present the results of impulse (numerical integration).

### 5.3.2. Statistical Results – Discrete Analysis

The mean and standard deviation of peaks or valleys (depending on the GRF component) for each group are presented in Table 10. All scalar extracted (anterior posterior first valley, vertical first peak, and medial-lateral first valley) presented a mean value smaller for the LAPAFI participants when compared with the participants of the GaitRec dataset groups (young, adult, and older adult).

*Table 10 Descriptive statistics GRF components 1st peak/valley*

| | **Anterior Posterior 1st Valley (% BW)** | | **Vertical 1st Peak (% BW)** | | **Medial-Lateral 1st valley (% BW)** | |
|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| **LAPAFI** | -0.13 | 0.03 | 0.99 | 0.05 | -0.019 | 0.007 |
| **GaitRec young** | -0.18 | 0.03 | 1.11 | 0.07 | -0.045 | 0.015 |
| **GaitRec adult** | -0.18 | 0.03 | 1.13 | 0.07 | -0.037 | 0.013 |
| **GaitRec older adult** | -0.17 | 0.04 | 1.08 | 0.08 | -0.044 | 0.022 |

SD: Standard Deviation. Peaks or valleys were extracted from mean curves between right and left foot for each GRF component. %BW: bodyweight percentage.

The mean and standard deviation of total impulse for each group is presented in Table 11. The participants from LAPAFI dataset presented lower total impulse for anterior posterior and vertical components of the GRF vector, and greater medial-

lateral total impulse when compared with the participants of the GaitRec dataset groups (young, adult, and older adult).

*Table 11 Descriptive statistics GRF components total Impulse*

| | Anterior Posterior Total Impulse | | Vertical Total Impulse | | Medial-Lateral Total Impulse | |
|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **mean** | **SD** | **Mean** | **SD** |
| **LAPAFI** | 6.47 | 1.0 | 78.4 | 2.4 | 3.6 | 1.1 |
| **GaitRec young** | 9.45 | 1.3 | 80.9 | 1.3 | 2.0 | 0.5 |
| **GaitRec adult** | 9.31 | 0.9 | 80.4 | 1.5 | 2.1 | 0.7 |
| **GaitRec older adult** | 8.88 | 1.5 | 79.8 | 1.9 | 2.1 | 0.8 |

SD: Standard Deviation. Total impulses are calculated using mean curve between right and left foot for each GRF component. Impulses are calculated using numerical integration of absolute values of GRF normalized on amplitude (% bodyweight) and on time (% stance phase). Normalized Impulse Unit: ($\%BW \cdot \%Stance\_Phase$)

The Hotelling's $T^2$ test revealed that there are statistically significant differences between the LAPAFI and all groups of the GaitRec dataset for both peaks/valleys and total impulses, as presented on Table 12. Furthermore, the test failed to find statistically significant differences within the GaitRec groups. For those groups where the null hypothesis has been rejected (H0 reject = 1), a post-hoc test is necessary to account the influence of each GRF component.

Table 12 Hotelling's results $T^2$ with alpha=0.05 - 1st peak/valley and total impulse of GRF components

| df = (3,22) | Hotelling's $T^2$ Peaks | | | Hotelling's $T^2$ Impulses | | |
|---|---|---|---|---|---|---|
| | **Z score** | **P value** | **H0 reject** | **Z score** | **P value** | **H0 reject** |
| **GaitRec: young vs adult** | 2.52 | 0.5 | 0 | 0.955 | 0.8 | 0 |
| **GaitRec: young vs older adult** | 0.927 | 0.8 | 0 | 2.93 | 0.5 | 0 |
| **GaitRec: adult vs older adult** | 6.55 | 0.1 | 0 | 0.866 | 0.9 | 0 |
| **LAPAFI vs GaitRec Young** | 44.8 | <0.001 | 1 | 88.5 | <0.001 | 1 |
| **LAPAFI vs GaitRec adult** | 38.9 | <0.001 | 1 | 103 | <0.001 | 1 |
| **LAPAFI vs GaitRec older adult** | 15.7 | 0.01 | 1 | 47.2 | <0.001 | 1 |

SD: standard deviation. Z score: position of a data related to the mean. P value: estimated probability of rejecting the null hypothesis. H0 reject: to reject (1) or not (0) the null hypothesis that the means are not equal. The multivariable comparison of peak/valley is related to the results from Table 10, and for impulses from Table 11.

The results of the post-hoc t-test (with Šidák p-value correction for three variables: p=0.0170) applied for each component of the GRF vector for peaks/valleys is presented on Table 13 and for total impulses on Table 14.

*Table 13 Post hoc t-tests GRF components 1st peak/valley with p_critical = 0.0170 (Šidák corrected)*

| | **t-test AP peak** | | | **t-test V peak** | | | **t-test ML peak** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Z score** | **P value** | **H0 reject** | **Z score** | **P value** | **H0 reject** | **Z score** | **P value** | **H0 reject** |
| **LAPAFI vs GaitRec Young** | 4.21 | <0.001 | 1 | -4.55 | <0.001 | 1 | 5.22 | <0.001 | 1 |
| **LAPAFI vs GaitRec adult** | 4.49 | <0.001 | 1 | -5.36 | <0.001 | 1 | 4.04 | <0.001 | 1 |
| **LAPAFI vs GaitRec older adult** | 2.78 | 0.01 | 1 | -3.10 | 0.005 | 1 | 3.63 | <0.001 | 1 |

SD: standard deviation. Z score: position of a data related to the mean. P value: estimated probability of rejecting the null hypothesis. H0 reject: to reject (1) or not (0) the null hypothesis that the means are not equal.

*Table 14 Post hoc t-tests GRF components total impulse with p_critical = 0.0170 (Šidák corrected)*

| | **t-test AP Total Impulse** | | | **t-test V Total Impulse** | | | **t-test ML Total Impulse** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Z score** | **P value** | **H0 reject** | **Z score** | **P value** | **H0 reject** | **Z score** | **P value** | **H0 reject** |
| **LAPAFI vs GaitRec Young** | -6.09 | <0.001 | 1 | -3.06 | 0.006 | 1 | 4.75 | <0.001 | 1 |
| **LAPAFI vs GaitRec adult** | -6.95 | <0.001 | 1 | -2.38 | 0.03 | 0 | 4.19 | <0.001 | 1 |
| **LAPAFI vs GaitRec older adult** | -4.48 | <0.001 | 1 | -1.57 | 0.1 | 0 | 3.91 | <0.001 | 1 |

SD: standard deviation. Z score: position of a data related to the mean. P value: estimated probability of rejecting the null hypothesis. H0 reject: to reject (1) or not (0) the null hypothesis that the means are not equal.

Regarding the peaks/valleys, we observed that all components contribute significantly to reject the null hypothesis for all groups' comparison. Regarding the total impulses, only the vertical component for LAPAFI compared with the adult and older adult groups of GaitRec dataset did not show statistically significant contribution to reject the null hypothesis. In the next section we produce comparisons, using the scalar fields instead of scalar points, then we will discuss if there is difference on results depending on the method of analysis.

## 5.4. **Statistical Parametric Mapping - SPM1D**

### 5.4.1. Independent Hotelling's $T^2$ test

a) GaitRec (within dataset groups):

The SPM vector field analysis of GRF components did not find regions with probability (p) values that indicate the likelihood with each suprathreshold cluster is expected to have been produced by a random field process with the same temporal smoothness for the entire stance phase (Figure 27, Figure 28, and Figure 29). Thus, post-hoc t-tests were not necessary. It is coherent with the findings from discrete scalar extracted on the traditional analysis method for the same group comparison combinations.

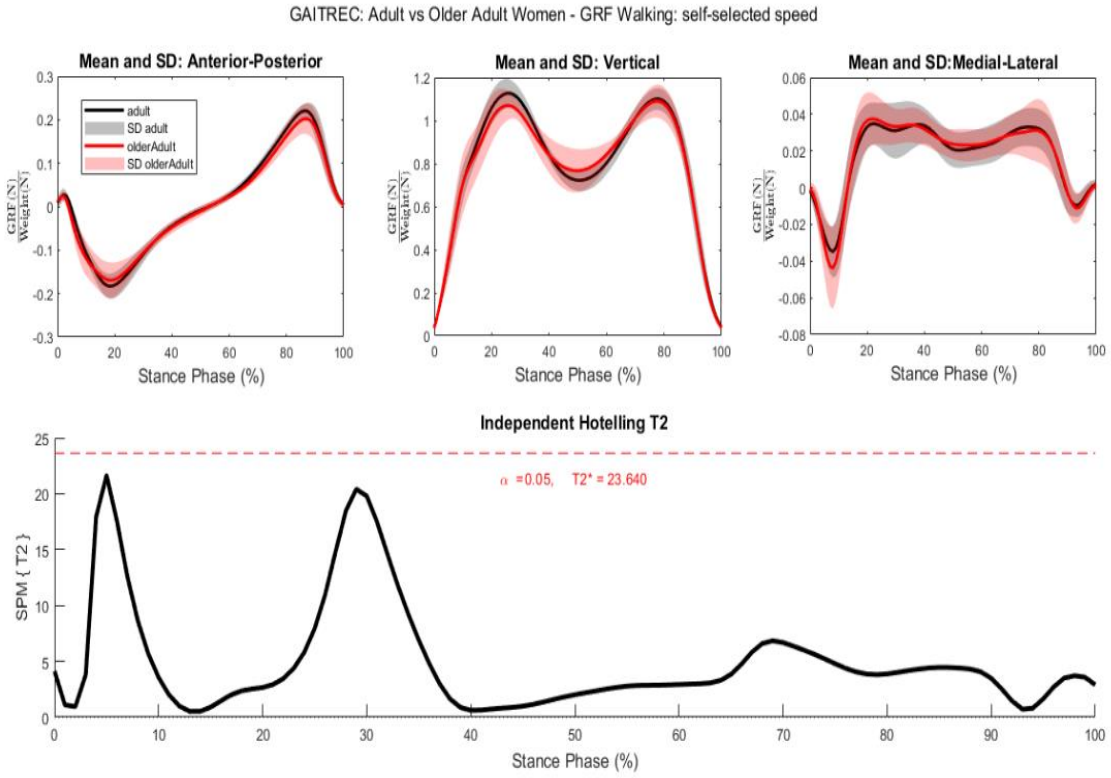



*Figure 27 Independent Hotelling T²: GaitRec young vs adult GRF scalar fields (The Author, 2021)*

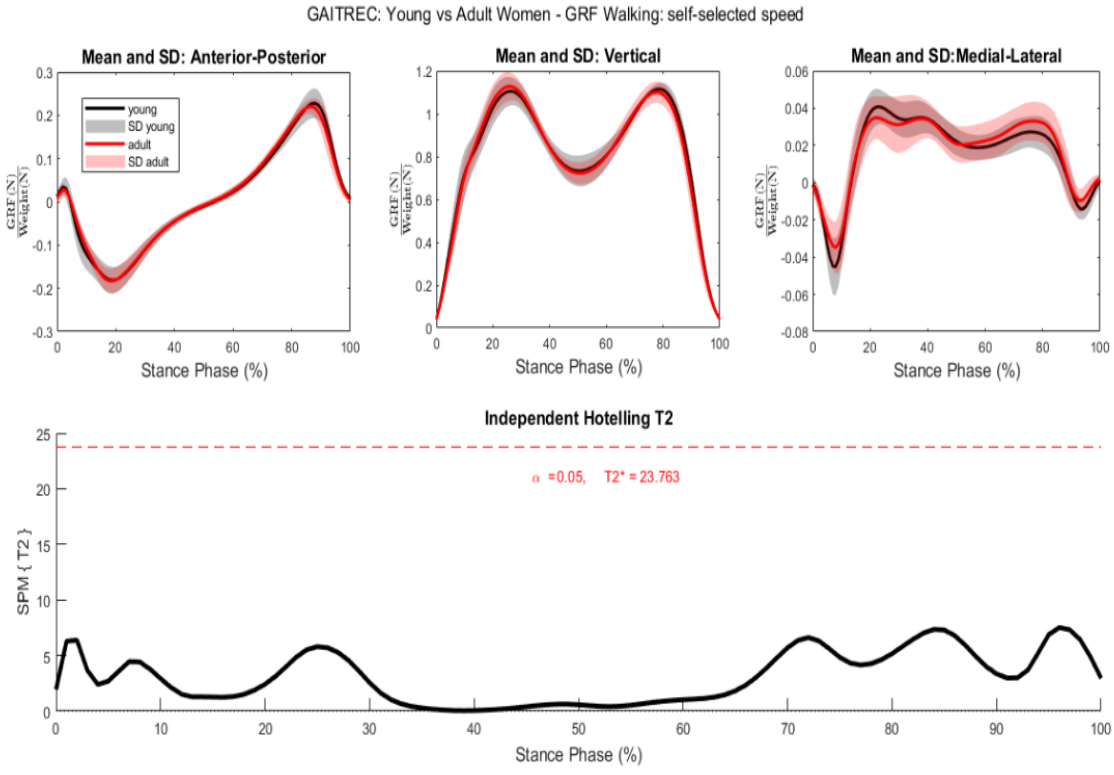



*Figure 28 Independent Hotelling T²: GaitRec young vs older adult GRF scalar fields (The Author, 2021).*

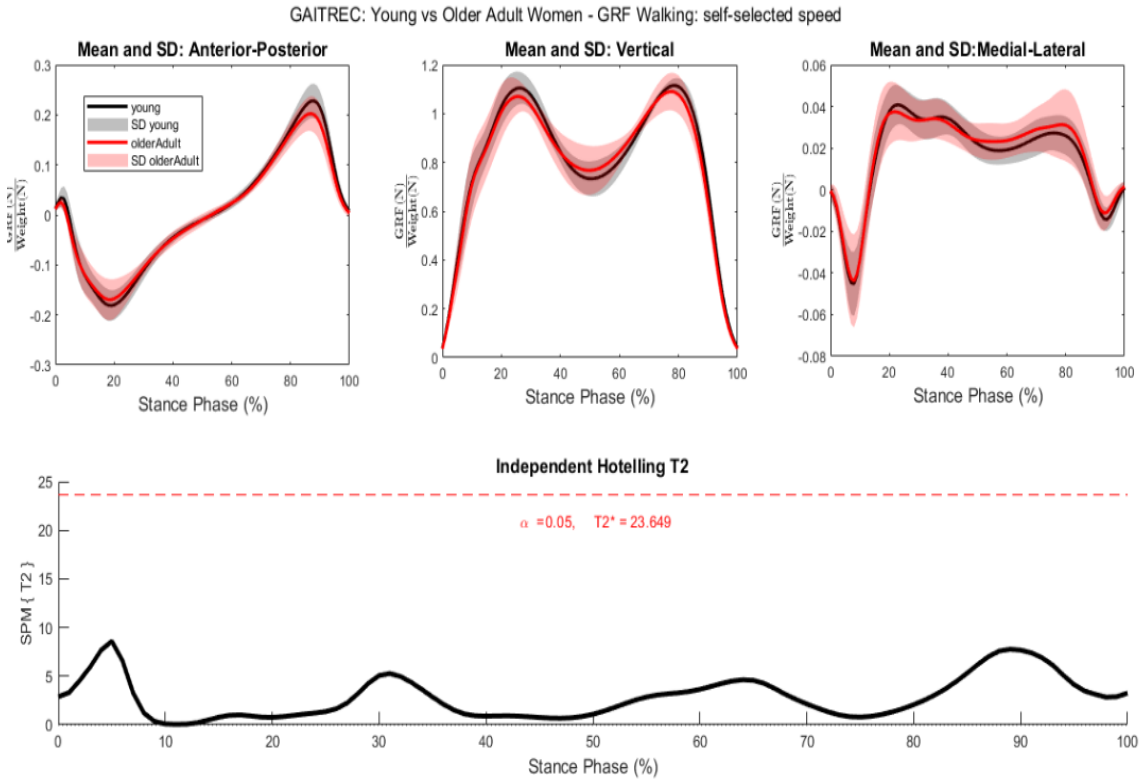



*Figure 29 Independent Hotelling T²: GaitRec adult vs older adult GRF scalar fields (The Author, 2021).*

b) GaitRec young vs LAPAFI older adults:

The SPM vector field analysis SPM{T²} between the young participants from GaitRec dataset and the older adults participants of the LAPAFI dataset found clusters surpassing the threshold T²* = 23.301. These were observed at approximately (% stance, specific p-value): (1%, p=0.049), (4-12%, p=0.001), (17-39%, p<0.001), and (43-100%, p<0.001), as presented numerically on Table 15 and graphically on Figure 30. The post-hoc analysis found significant contributions from all three GRF components. Anterior posterior: the GRF on this component was greater for young participants, surpassing a t*=4.054, over two clusters at approximately: (17-34%, p<0.001) and (64-94%, p<0.001). Vertical: the GRF on this component was greater for young participants, surpassing a t*=4.036, over two clusters at approximately: (23-32%, p<0.001) and (74-89%, p<0.001), and smaller over two clusters at approximately: (1%, p=0.017), (51-63%, p<0.001), and (97-100%, p=0.012). Medio lateral: the GRF on this component was greater for young participants, surpassing a t*=4.047, over one cluster at approximately: (5-10%, p<0.001), and smaller over two clusters at

approximately: (29-32%, p=0.006), and (46-88%, p<0.001). The post-hoc results for these groups are presented graphically on Figure 30 and numerically on Table 16.

c) GaitRec adult vs LAPAFI older adults:

The SPM vector field analysis SPM{$T^2$} between the older adult participants from GaitRec dataset and the adult participants of the LAPAFI dataset found clusters surpassing the threshold $T^{2*}$ = 23.349. These were observed at approximately (% stance, specific p-value): (1%, p=0.049), (3-5%, p=0.039), (8-11%, p<0.03), (16-37%, p<0.001), and (44-100%, p<0.001) as presented graphically on Figure 31 and numerically on Table 15. The post-hoc analysis found significant contributions from all three GRF components. Anterior posterior: the GRF on this component was greater for adult participants, surpassing a t*=4.070, over three clusters at approximately: (2%, p=0.017), (16-34%, p<0.001), and (62-91%, p<0.001). Vertical: the GRF on this component was greater for adult participants, surpassing a t*=4.033, over two clusters at approximately: (23-32%, p<0.001) and (74-89%, p<0.001), and smaller over three clusters at approximately: (1%, p=0.016), (45-62%, p<0.001), and (98-100%, p=0.013). Medio lateral: the GRF on this component was greater for young participants, surpassing a t*=4.047, over one cluster at approximately: (6-9%, p=0.011), and smaller over one clusters at approximately: (46-72%, p<0.001). The post-hoc results for these groups are illustrated on Figure 31 and numerically on Table 16.

d) GaitRec older adult vs LAPAFI older adult:

The SPM vector field analyses (SPM{$T^2$}) between the older adult participants from GaitRec dataset and the older adult participants of the LAPAFI dataset found clusters surpassing the threshold $T^{2*}$ = 23.349. These were observed at approximately (% stance, specific p-value): (1%, p=0.049), (20-40%, p<0.001), and (56-100%, p<0.001).as presented graphically on Figure 32 and numerically on Table 15. The post-hoc analysis found significant contributions from all three GRF components. Anterior posterior: the GRF on this component was greater for older adults (GaitRec) participants, surpassing a t*=4.077, over two clusters at approximately (% stance, specific p-value): (16-34%, p<0.001), and (62-91%, p<0.001). Vertical: the GRF on this component was smaller for older adult (GaitRec) participants, surpassing a t*=4.024, over two clusters at approximately: (1%, p<0.001) and (96-100%, p=0.006). Medial-

lateral: the GRF on this component was greater for young participants, surpassing a t*=4.022, over one cluster at approximately: (7-9%, p=0.013), and smaller over two clusters at approximately: (27-33%, p<0.001) and (44-73%, p<0.001). The post-hoc results for these groups are illustrated on Figure 32 and numerically on Table 16.

These results will be explored over the next section. First, by pointing out the main observations from data preparation processes. In addition, we discuss the outcomes from the comparison between traditional discrete analysis and the SPM method, as well as their usability. Finally, we will discuss about the validity of the results from the comparisons between groups (divided by age) confronting with physiological aspects evidence from literature. Further, we present the main limitations and suggestions for future work.

*Table 15 Summary results of GRF components: independent Hotelling's $T^2$.*

| SPM(HOBBS; ROBINSON; CLAYTON) | | GaitRec vs LAPAFI | | |
|---|---|---|---|---|
| | | **Young vs LAPAFI** | **Adult vs LAPAFI** | **Older adult vs LAPAFI** |
| **$T^2$*** | | 23.30 | 23.35 | 23.24 |
| **cluster 1** | **% Stance** | 1 | 1 | 1 |
| | **p-value** | 0.049 | 0.049 | 0.049 |
| **cluster 2** | **% Stance** | 4-12 | 3-5 | 20-40 |
| | **p-value** | 0.001 | 0.039 | <0.001 |
| **cluster 3** | **% Stance** | 17-39 | 8-11 | 56-100 |
| | **p-value** | <0.001 | 0.03 | <0.001 |
| **cluster 4** | **% Stance** | 43-100 | 16-37 | - |
| | **p-value** | <0.001 | <0.001 | - |
| **cluster 5** | **% Stance** | - | 44-100 | - |
| | **p-value** | - | <0.001 | - |

$T^2$*: Critical threshold for rejecting the null hypothesis. % Stance: percentual (instant or range) of gait stance phase where there is significant difference. p-value: estimated probability of rejecting the null hypothesis.

*Table 16 Summary of post-hoc t-test results of individual GRF components*

| Post hoc t test Sidák | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **GRF Anterior posterior** | | | **cluster 1** | | **cluster 2** | | **cluster 3** | | | |
| **GaitRec vs LAPAFI** | **t*** | **% stance** | **p-value** | **% stance** | **p-value** | **% stance** | **p-value** | | | |
| Young vs LAPAFI | 4.054 | 17-34 | <0.001 | 64-94 | <0.001 | - | - | | | |
| Adults vs LAPAFI | 4.070 | 2 | 0.017 | 16-34 | <0.001 | 62-91 | <0.001 | | | |
| Older adults vs LAPAFI | 4.077 | 24-34 | <0.001 | 66-90 | <0.001 | - | - | | | |
| **GRF Vertical** | | **cluster 1** | | **cluster 2** | | **cluster 3** | | **cluster 4** | | **cluster 5** | |
| **GaitRec vs LAPAFI** | **t*** | **% stance** | **p-value** | **% stance** | **p-value** | **% stance** | **p-value** | **% stance** | **p-value** | **% stance** | **p-value** |
| Young vs LAPAFI | 4.036 | 1 | 0.017 | 23-32 | <0.001 | 51-63 | <0.001 | 74-89 | <0.001 | 97-100 | 0.012 |
| Adults vs LAPAFI | 4.033 | 1 | 0.016 | 20-34 | <0.001 | 45-62 | <0.001 | 78-84 | 0.003 | 98-100 | 0.013 |
| Older adults vs LAPAFI | 4.024 | 1 | 0.016 | 96-100 | 0.006 | - | - | - | - | - | - |
| **GRF Medial-lateral** | | | **cluster 1** | | **cluster 2** | | **cluster 3** | | | |
| **GaitRec vs LAPAFI** | **t*** | **% stance** | **p-value** | **% stance** | **p-value** | **% stance** | **p-value** | | | |
| Young vs LAPAFI | 4.047 | 5-10 | 0.002 | 29-32 | 0.006 | 46-88 | <0.001 | | | |
| Adult vs LAPAFI | 4.056 | 6-9 | 0.011 | 46-72 | <0.001 | - | - | | | |
| Older adult vs LAPAFI | 4.022 | 7-9 | 0.013 | 27-33 | 0.002 | 44-73 | <0.001 | | | |

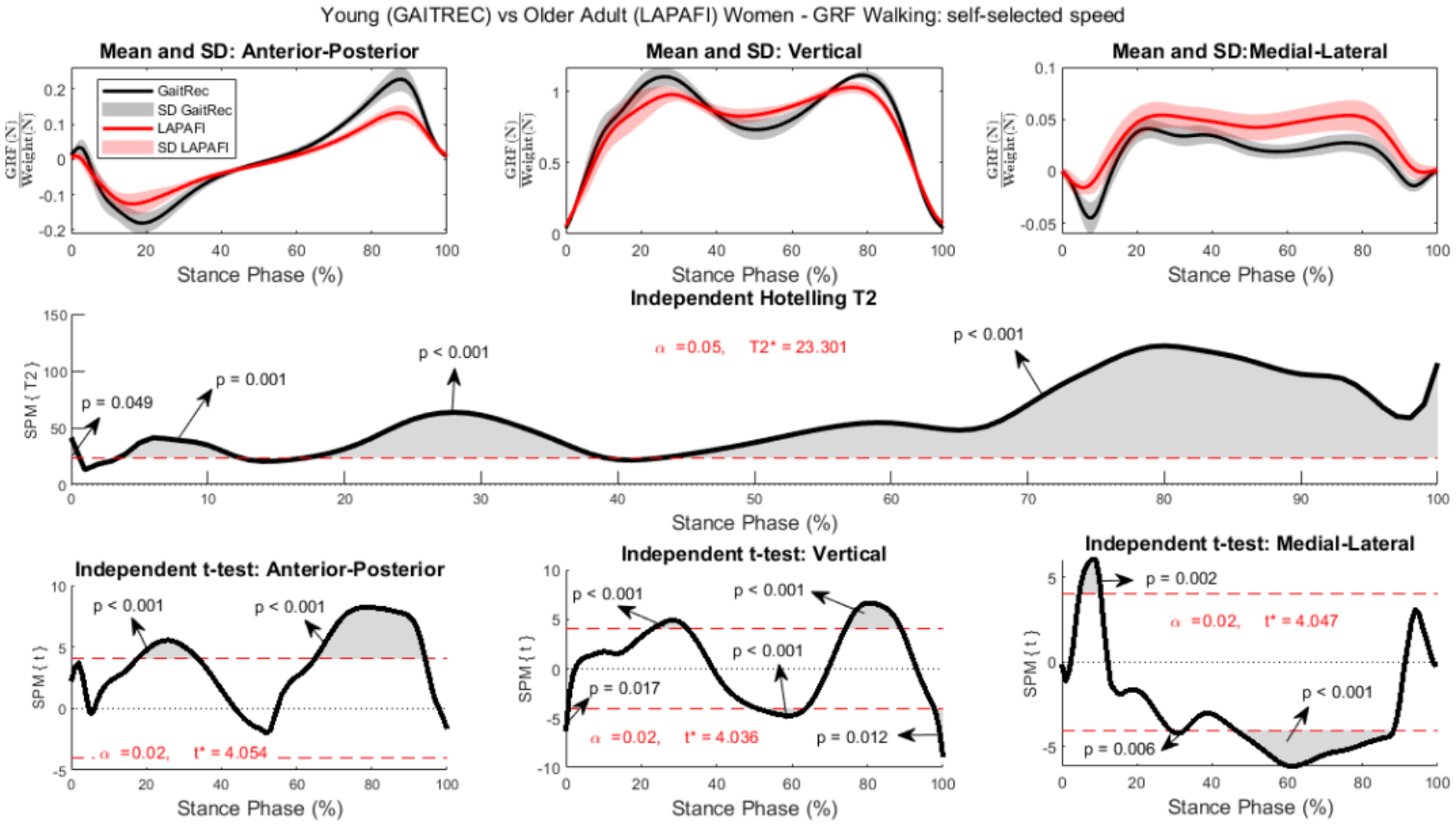



*Figure 30 Independent Hotelling T² GaitRec (young) vs LAPAFI (older adults) GRF scalar fields. Post-hoc scalar field GRF independent t-test p-critical = 0.0170 (The Author, 2021).*

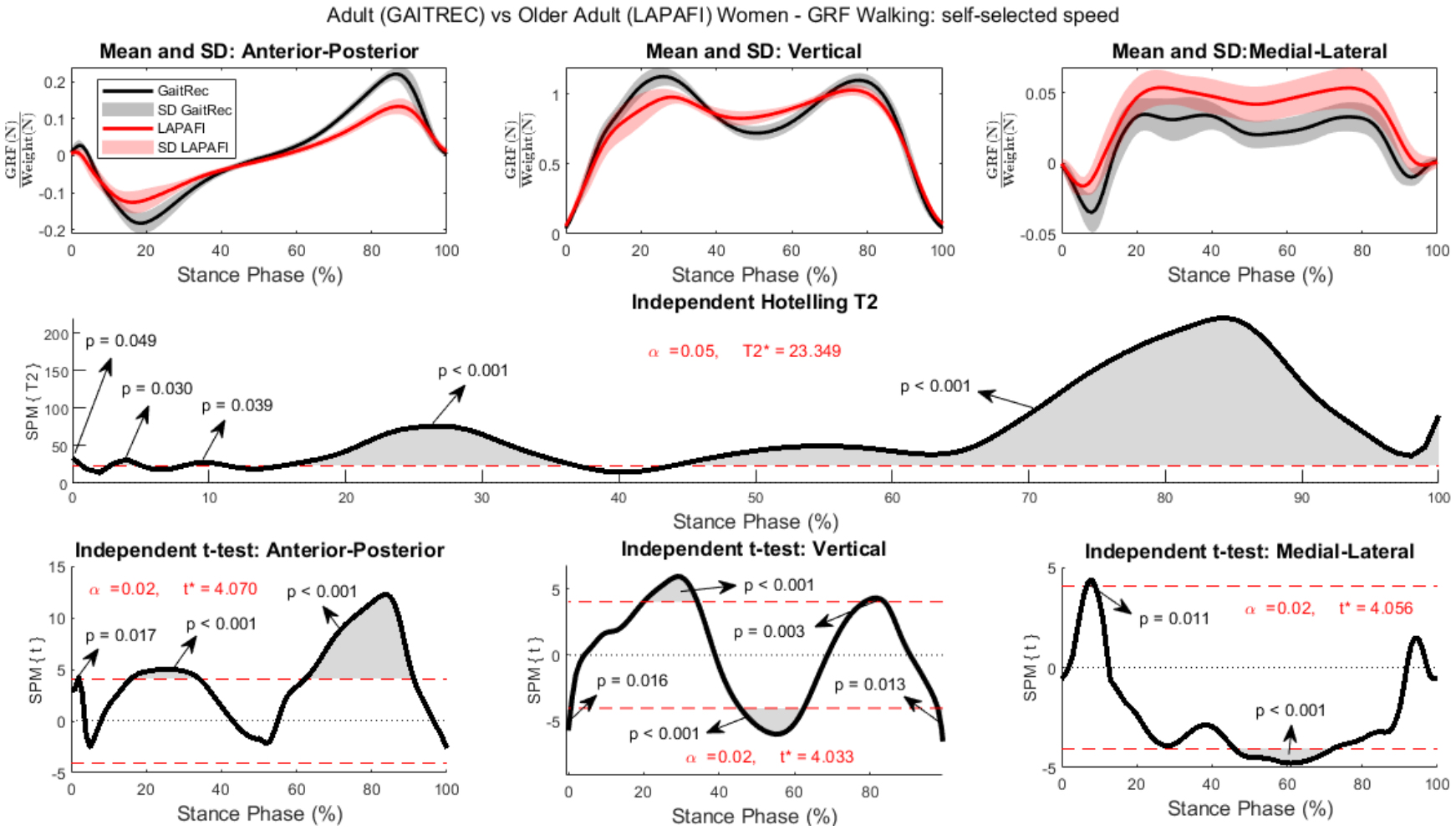



*Figure 31 Independent Hotelling T² GaitRec (adults) vs LAPAFI (older adults) GRF scalar fields. Post-hoc scalar field GRF independent t-test p-critical = 0.017 (The Author, 2021).*

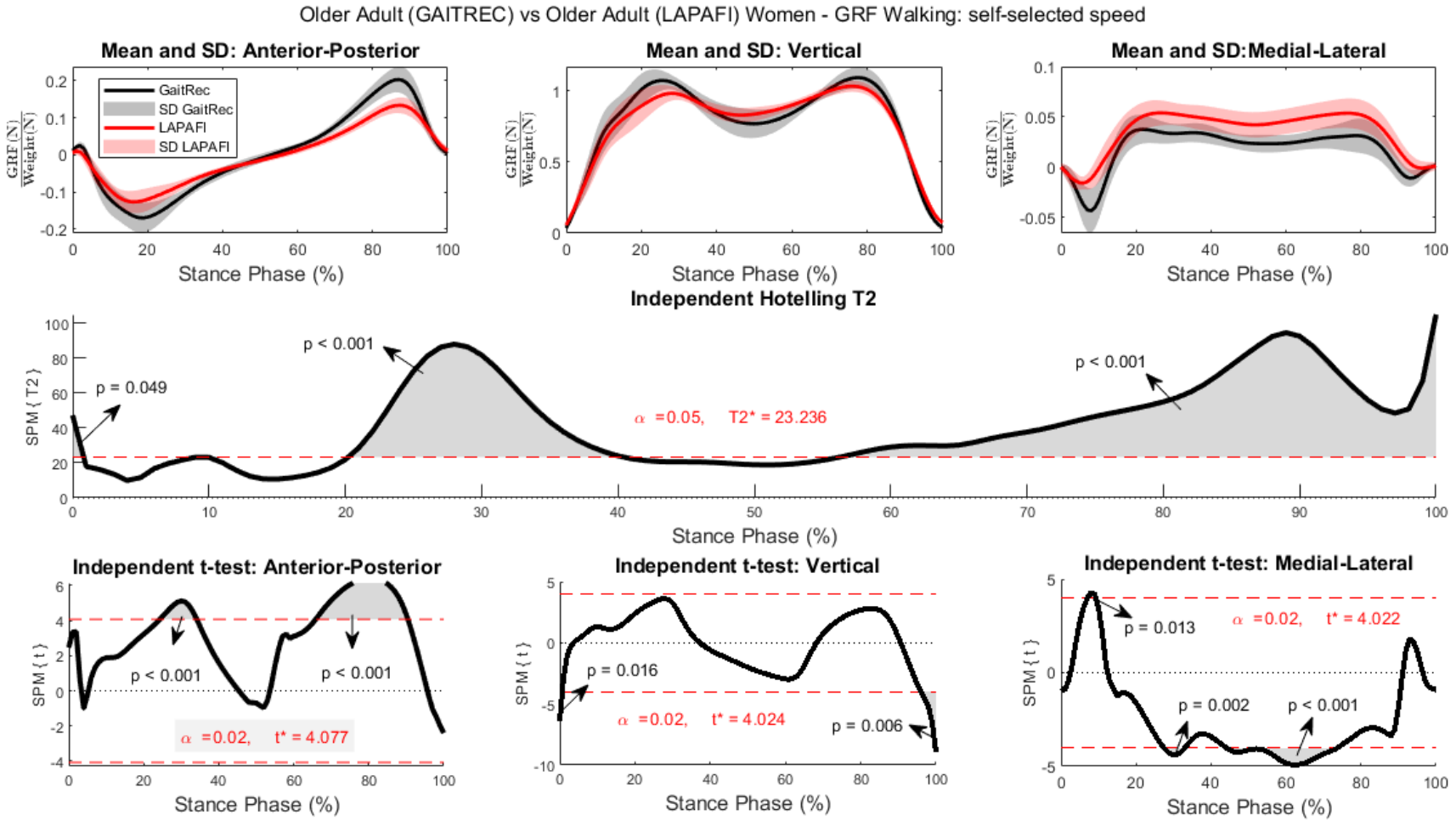



*Figure 32 Independent Hotelling T²: GaitRec (older adults) vs LAPAFI (older adults) GRF scalar fields. Post-hoc scalar field GRF independent t-test p-critical = 0.0170 (The Author, 2021).*

# 6. DISCUSSION

## 6.1. GRF Data Preparation

The poor data quality is a prevalent aspect of large datasets and on the internet, therefore it is a well-known problem on data science applications (SAHA; SRIVASTAVA, 2014). There are stablishing practical questions to improve data quality (MCCORD; WELTY; COURTWRIGHT; DILLON *et al.*, 2021). As we will discuss in the following section, preliminary processes on data are important to improve the reliability on the outcomes of analysis.

Regarding the intrasubject correlation filter, the procedure of removing discrepant curves has improved the quality of the data accepted for final analysis with the cost of losing a participant. The method seems effective for participants with a greater number of acceptable curves. However, it is not as effective for participants with a small number of acceptable curves, as presented for participant ID=2 on the results section.

For further validation, the method for data quality assurance should be confronted with other methods of similar objective, which was not the general purpose of this study, but a preliminary step to apply the analysis methods. Moreover, the intersubjects correlation filter was important, especially because we did not have the information on the speed of each participant. However, it costs the loss of two subjects, making our sample even smaller, which decreases the intragroup variability and does not contribute to SPM results reliability.

## 6.2. GRF Discrete and Continuous Analysis

Regarding the traditional discrete analysis and the continuous analysis method SPM, the outcomes from both methods were consistent. It is, both methods were able to reject the null-hypothesis on their domain of search (0D for scalar extraction and 1D for vector fields). However, the SPM1D method was able to expand the analysis showing regions that rejected the null hypothesis which would not be explored if only the discrete analysis were performed. This result contrasts with some authors that have observed ambiguous conclusions from the traditional approach and defend that the SPM method avoids this problem (PATAKY; VANRENTERGHEM; ROBINSON, 2015). It is because just one test leads to a single conclusion, instead of having to choose between multiple options of discrete parameters (specific time point or spatial location).

Justified by the fact that vector field testing overcomes the most common bias sources: failure to consider the entire domain and failure to consider the covariance amongst vector components (PATAKY; ROBINSON; VANRENTERGHEM, 2013).

Furthermore, the authors claim that scalar field SPM solves the regional focus bias by testing the entire domain $q$, but they also warn that the method remains susceptible to covariance bias because it separately tests the $I$ vector components. They also warn that non-RFT corrections like the Šidák, which was applied on this research, can partially solve the problem of multiple comparisons, because they fail to consider the spatiotemporal smoothness of the measurement domain $q$, overestimating the number of independent tests. It leads to an overly conservative threshold except for very rough fields. Besides that, non-RFT corrections also fail to solve covariance bias because they assume that vector components vary independently. They also acknowledge that there are many additional sources of bias which are not exclusive to SPM and that misregistration of trajectories and unit normalization are some examples of common problems for scalar extraction and vector field (PATAKY; ROBINSON; VANRENTERGHEM, 2013).

Regarding the datasets, the results from discrete analysis and SPM1D were consistent. The GaitRec scalar values extracted (force peak/valley and impulses) did not show statistical differences when compared between age groups (GaitRec young vs adult, young vs older adult, and adult vs older adult). Similarly, the same was observed for the continuous analysis: no cluster achieved the critical threshold when compared between age groups as previously declared. The vector field SPM analysis found differences for all GaitRec groups when compared to LAPAFI participants on almost all time-series. These results should be interpreted with caution as the comparison between groups of the same laboratory (GaitRec) failed to find clusters that trespassed the critical threshold ($T^{2*}$).

The origin of these differences may be due to: (i) speed control (measurement), although all groups walked at self-selected speed, and we do not know the speed the LAPAFI participants chose as self-selected. (ii) age: LAPAFI participants ($67 \pm 7\ years$) are older than all groups of GaitRec (young ($35 \pm 4\ years$), adult ($40 \pm 6\ years$), older adult ($54 \pm 4.5\ years$)), as presented on characterization of participants on the results section. (iii) Experimental aspects: type of force platforms (type of sensors, affects sensibility), parameters setting of acquisition (force threshold

differences), procedures of data processing (sampling, interpolation method, filtering methods).

Regarding the sensibility of time-continuous analysis (SPM) to timing variability of gait data, there is evidence that gait interventions, such as changes in walking speed, can induce temporal shifts that affect time-continuous results from both intrasubject and intersubjects analysis perspectives, encouraging researchers to investigate how both amplitude and timing of biomechanical data affects the time continuous analysis results (HONERT; PATAKY, 2021). To overcome the alignment problem of data curves, there are solutions being developed such as the manual one-dimensional data warping and nonlinear registration (mwarp1d) in Python and PyQt, which provides a collection of graphical user interface (GUI) and scripting tools to do the manual registration of 1D datasets, thereby achieving qualitatively optimal alignment of multiple 1D observations (PATAKY; NAOUMA; DONNELLY, 2019).

It is worthy to mention that, due to methodological limitations on detecting the gait events on medio-lateral GRF vector component, Appendix G and Appendix H, we could not use other discrete points as parameters to test other time windows, which could confirm the narrative built from literature evidence. So, although the SPM overcomes the main sources of bias, it does not replace the necessity of identifying those discrete parameters. Furthermore, those can be important allies to overcome other bias sources such as misregistration of curves, and units normalization (amplitude and timing).

The main disadvantage observed on applying the SPM on GRF experimental data was not the complexity involved to understand and develop the SPM scripts, those are well documented with examples for different datasets, and experimental designs, both synonyms on how data is mounted or called by script. It can be confusing to deal with in a first contact, but undoubtedly, the dependency between the reliability of the SPM results on data preparation, and the specific aspects of the datasets available for analysis, was the most time-consuming task, and is referred as preliminaries on SPM documentation.

The main advantage of applying SPM on biomechanics is the possibility of analyzing the whole time-series in one test, even for multiple variables (SPM{$T^2$}) and multiple groups (ANOVA), in an intrasubject and intersubjects perspective. It is possible as long as we can rely on our data, and the processes behind the whole cycle of planning, training people, acquiring, storing, sharing, managing and transforming

data into important information through multidisciplinary analysis and communication, on consonance with practitioners (physicians, physiotherapists, physical educators) and technological agents.

### 6.3. **GRF Age Groups Comparison**

From the traditional discrete analysis results, there appears to be an inverse relationship between age and anterior posterior and vertical total impulse, it means that older participants seem to present lower total impulse for the anterior posterior and vertical components of the GRF vector. However, on the medial-lateral the total impulse seems to increase with the advance of age. There was also a decreased force first peak/valley for all GRF components with aging, with exception of the young group (GaitRec) for anterior-posterior and medio-lateral GRF components when compared with the adult group (GaitRec).

For anterior-posterior GRF component, lower first peak (deceleration phase) was observed for older adults when compared with younger adults (LARISH; MARTIN; MUNGIOLE, 1988). The scalar extraction results agree with literature observations. However, the hypothesis tests did not find statistically significant differences between the GaitRec groups, but it did find differences between the LAPAFI older adult participants and all groups of the GaitRec (young, adults and older adults) for both scalar extraction and continuous (SPM) methods of analysis.

The decrease on force for older adults was expected as aging effects the GRF components. The literature has observed algebraic difference for the vertical component force peaks, which suggests there is an age-related decrease in vertical oscillations of the center of gravity, resulting in a lower vertical acceleration in the center of gravity for the older adults, which could be interpreted as an attempt by older adults to improve economy of energy expenditure (LARISH; MARTIN; MUNGIOLE, 1988).

A study on age and gender comparisons of muscle strength with 654 participants (men and women), shows evidence that knee extensors muscles (rectus femoris (hip flexor) and vasti group) quality (strength per kilogram of regional free fat mass (FFM)) declines with age in both men and women. The age associated with concentric (Con) strength losses seems to begin at age 40 for both genders, while the eccentric (Ecc) strength losses appear to happen later on women. The authors suggest that older women have a greater ability to store and utilize elastic energy via the

stretch-shortening cycle (from Ecc action), even when compared with younger participants (LINDLE; METTER; LYNCH; FLEG *et al.*, 1997).

Using simulation of walking to examine muscle contributions on forward progression (AP), and support (V) of the body mass center was possible to identify that the first half of stance phase is mostly managed by the vasti group and gluteus maximus on AP deceleration and V acceleration with contributions from dorsiflexors and Soleus. In the single limb support (32%-50% stance phase) the gluteus medius accelerates the body forward and upwards. In the second half of stance, forward and upward acceleration are mainly produced by gastrocnemius and soleus, with addition of some uniarticular plantarflexors for the upward direction (LIU; ANDERSON; PANDY; DELP, 2006).

The LAPAFI participants (older adults) when compared to young and adults from GaitRec: the multivariate SPM{$T^2$} test led to clusters' observation over almost the entire time-series. Post-hoc t-tests revealed that the anterior-posterior component of GRF vector contributes to the differences on loading response and over terminal and pre-swing sub-phases of stance, presenting lower (module) values for the older adult (LAPAFI) participants.

The observation of those clusters may be related to differences on the vasti group and gluteus maximus during early stance. For the second half of stance those clusters can be related to gastrocnemius and soleus, with addition of some plantarflexors. The vertical GRF vector component contributes to the observation of clusters on both extrema: heel strike and toe off, presenting greater values on these regions for the older adults. The older adults also presented greater force over mid stance and smaller force over loading response and over the transition of terminal stance and pre-swing subphases of stance.

So, older adults appear to rely more on the support over mid-stance than younger participants, but less on the deceleration and acceleration phases of stance, it may be related to those muscle quality aspects mentioned above, so we can conjecture that older adults are more confident during stable events (midstance) than during unstable events (peaks on first and second half of stance).

The medial-lateral component of GRF vector contributes to the observation of clusters, presenting lower values (module) at the begin of load response and greater values at the beginning of midstance and over terminal stance. Here, we can

conjecture that the weaker force on heal-strike is a defense mechanism to avoid a medio lateral instability before building the single limb stance.

The LAPAFI participants when compared to older adults from GaitRec: the multivariate SPM{$T^2$} test led to observation of clusters over initial contact, between transition of loading response and mid-stance and over half of terminal stance to the entire pre-swing subphases of stance. Post-hoc t-tests revealed that the anterior-posterior component of GRF vector contributes to differences on loading response and over terminal and pre-swing sub-phases of stance, presenting lower (module) values for the older adult (LAPAFI) participants.

The vertical GRF vector component contributes to differences on both extrema: heel strike and toe off, presenting greater values on these regions. The medial-lateral component of GRF vector contributes presenting lower values (module) at the begin of load response and greater values at the beginning of midstance and over terminal stance. The fact that groups of closer age ranges did not show differences on vertical GRF, beside on the extrema points of stance phase, may be explained by the fact that both behave similarly: reducing force on the acceleration and deceleration, but maintaining on mid-stance.

The results from SPM{$T^2$} analysis between LAPAFI and GaitRec agree with the literature suggestions that elderly presents lower first and second peak and a higher minimum value at mid-stance (vertical GRF) which has not been observed by Toda (2015). On his observations, only the second peak of the vertical GRF of the elderly women is found to be lower compared with young women (TODA; NAGANO; LUO, 2015).

However, it is not confirmed by the analysis between groups of GaitRec dataset. The most accepted hypothesis is that older adults present lower motion and rely less on ankle. For the AP GRF component, was observed a pattern of the second half stance presenting wider clusters on SPM tests that may corroborate with the findings of greater contribution from plantar flexors, relying less on ankles (BOYER; JOHNSON; BANKS; JEWELL *et al.*, 2017).

The authors suggest that older women have a greater ability to store and utilize elastic energy via the stretch-shortening cycle (from Ecc action of knee extensor), even when compared with younger participants, which is observed considering the greater GRF during mid-stance, where the knee is in Ecc action (LINDLE; METTER; LYNCH; FLEG *et al.*, 1997).

## 6.4. **Limitations and Suggestions for Future Studies:**

Regarding the protocol of data acquisition and management: the protocol of data acquisition can be improved making it less susceptible to human errors. Invest on creating a more robust and well documented methodology for data acquisition considering the following:

a) GRF, COP, Speed Control and Procedural Suggestions

Both GRF and COP should be exported and included in future analysis. Greater attention should be given on the quality of the data acquired to minimize the amount of data that is lost during data preparation. Especially when human intervention is needed to decide boundaries (such as informing the start and end events of a gait step). If possible, implement a systematic methodology for data storage and sharing is also suggested.

A specific methodology for speed measurement should be developed and implemented for speed control (even for self-selected speed). Greater attention should be given to the process of exporting data. The suggestion is to invest on developing techniques for automatization of this process, because it currently relies on the technician eyesight to specify when a step starts and when it ends. Even a small deviation could at some level compromise the data.

The GRF threshold of force platforms should be set to a lower value than 30 N, because a lower value is possible as informed by the technical support of the manufacturer.

b) Anthropometric Measurements, Demographic Data and Experimental Conditions

Additional information on anthropometric features can be included, such as body segment lengths, and shoe/foot size. Demographic data could also be expanded and methodologically improved, investing on systematization. I would suggest implementing a digital form with mandatory fields, filled by a well-trained technician/researcher with questions like date of birth, date of examination and age. Health history related questions should be collected, such as: medications, historical conditions, previous surgeries, even for healthy participants. Physical activities habits,

and shoes and clothing condition during examination can also be included on that digital form.

The digital form can be developed based on the research question, but a set of default questions can be defined by the laboratory to build a robust database that complements the measurements made with the vast equipment available. I reinforce the importance of stablishing a systematization of each step of data acquisition which may lead to improvement of data collection, and mitigation of data loss. Improvements on data storage, data management, and data processing are also possible, considering the current importance of securing data privacy policies, which is a whole branch of work that should be developed.

The main limitation of applying the SPM on biomechanical data are related to pre-processing which is time consuming and the lack of a graphical interface, which would improve the usability. The fact that SPM involves more complex calculations may prevent researchers with no prior knowledge to explore the tool, a graphical interface may not solve this issue, but it does encourage researchers to attempt to understand and apply the method on their studies. Regarding the test chosen for our study, we suggest running an ANOVA (analysis of variance) to compare between groups for each GRF component.

# 7. CONCLUSIONS

Regarding the main objective of this research, which was to explore the usability and main aspects of SPM applied to experimental GRF gait data, we conclude that beside the necessity of developing graphical tools to better operate the processes of time-continuous statistical testing, the main challenge involves preparing data. We have shown a correlation filter as an option to solve the curve cleaning problem which presented better efficiency for those participants with greater number of acceptable curves.

In summary, both methods of analysis were able to find differences, but the SPM method expands the search and leads to a more complete analysis. The importance of data preparation remains accountable for both methods, and the classical discrete method may be applied to overcome time-event related sources of bias. So, the methods are complementary to each other. Considering the datasets, there was no differences when compared between GaitRec groups, but differences were observed for all groups when compared with the LAPAFI datasets. These appeared mainly on extrema (begin and end) and optimal (maximum or minimum) GRF values of the stance phase. In conclusion, older adults from LAPAFI dataset when compared with the young and adult groups of GaitRec dataset, showed less strength on both deceleration and acceleration phases of gait, and greater strength on midstance. There was no differences observed on these time-windows when the LAPAFI dataset was compared with the older adults from GaitRec, only on the start and beginning. It bring us to these older adults women rely more on stabilizers, which are activated during the stance phase (simple, between first and second GRF peaks and containing mid stance).

Furthermore, when analyzing the outcomes of the SPM analysis in concomitance with the discrete gait events and their physiological connections, it is possible to expand from data to real world meaningful information and application. We reinforce the importance of data quality for a reliable analysis.

**REFERENCES**

BAKER, R. **Measuring Walking: A Handbook of Clinical Gait Analysis**. 2013. 229 p. ISBN-13: 978-1908316660.

BEGG, R. K.; WYTCH, R.; MAJOR, R. E. Instrumentation used in clinical gait studies: A review. **Journal of Medical Engineering & Technology**, 13, n. 6, p. 290-295, 1989/01/01 1989.

BOYER, K. A.; JOHNSON, R. T.; BANKS, J. J.; JEWELL, C. *et al.* Systematic review and meta-analysis of gait mechanics in young and older adults. **Experimental Gerontology**, 95, p. 63-70, 2017/09/01/ 2017.

CASTRO, M. P.; PATAKY, T. C.; SOLE, G.; VILAS-BOAS, J. P. Pooling sexes when assessing ground reaction forces during walking: Statistical Parametric Mapping versus traditional approach. **Journal of Biomechanics**, 48, n. 10, p. 2162-2165, Jul 2015.

CAVAGNA, G. A. Physiological aspects of legged terrestrial locomotion. **Cham, Switzerland: Springer**, 2017.

CHOCKALINGAM, N.; HEALY, A.; NEEDHAM, R. Interpreting Ground Reaction Forces in Gait. *In*: **Handbook of Human Motion**. Cham: Springer International Publishing, 2018. p. 609-623.

CRUZ-JIMENEZ, M. Normal Changes in Gait and Mobility Problems in the Elderly. **Physical medicine and rehabilitation clinics of North America**, 28, n. 4, p. 713-725, 2017/11// 2017.

EDWARDS, W. B.; DERRICK, T. R.; HAMILL, J. Time Series Analysis in Biomechanics. *In*: **Handbook of Human Motion**. Cham: Springer International Publishing, 2018. p. 349-371.

ENDERLE, J. **Introduction to biomedical engineering**. Academic press, 2012. 1272 p. 978-0-12-374979-6.

F. VAVERKA, M. E., Z. SVOBODA, M. JANURA. System of gait analysis based on ground reaction force assessment. **Acta Gymnica**, 45, p. 187–193, 2015.

HOBBS, S. J.; ROBINSON, M. A.; CLAYTON, H. M. A simple method of equine limb force vector analysis and its potential applications. **Peerj**, 6, Feb 2018.

HONERT, E. C.; PATAKY, T. C. Timing of gait events affects whole trajectory analyses: A statistical parametric mapping sensitivity analysis of lower limb biomechanics. **Journal of Biomechanics**, 119, p. 110329, 2021.

HORSAK, B.; SLIJEPCEVIC, D.; RABERGER, A.-M.; SCHWAB, C. *et al.* GaiTRec, a large-scale ground reaction force dataset of healthy and impaired gait. **Scientific data**, 7, n. 1, p. 143-143, 2020.

KIRTLEY, C. **Clinical gait analysis: theory and practice**. Elsevier Health Sciences, 2006. 0443100098.

LARISH, D. D.; MARTIN, P. E.; MUNGIOLE, M. Characteristic patterns of gait in the healthy old. **Annals of the New York Academy of Sciences**, 515, p. 18-32, 1988 1988.

LINDLE, R.; METTER, E.; LYNCH, N.; FLEG, J. *et al.* Age and gender comparisons of muscle strength in 654 women and men aged 20–93 yr. **Journal of applied physiology**, 83, n. 5, p. 1581-1587, 1997.

LIU, M. Q.; ANDERSON, F. C.; PANDY, M. G.; DELP, S. L. Muscles that support the body also modulate forward progression during walking. **Journal of biomechanics**, 39, n. 14, p. 2623-2630, 2006.

MAHAKI, M.; SOUZA, G.; MIMAR, R.; VIEIRA, M. F. The comparison of ground reaction forces and lower limb muscles correlation and activation time delay between forward and backward walking. **Gait & Posture**, 58, p. 380-385, Oct 2017.

MALANGA, G.; DELISA, J. A. Section One. **Gait Analysis In The Science Of Rehabilitation**, 2, p. 1, 1998.

MCCORD, S. E.; WELTY, J. L.; COURTWRIGHT, J.; DILLON, C. *et al.* Ten practical questions to improve data quality. **Rangelands**, 2021.

OKUNO, E.; FRATIN, L. **Biomechanics of the human body**. Springer, 2014. 1461485754.

OLIVE, D. J.; CHERNYK. **Robust multivariate analysis**. Springer, 2017. 3319682539.

PATAKY, T.; VANRENTERGHEM, J.; ROBINSON, M. The probability of false positives in zero-dimensional analyses of one-dimensional kinematic, force and EMG trajectories. **Journal of biomechanics**, 49 9, p. 1468-1476, 2016.

PATAKY, T. C. Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. **Journal of biomechanics**, 43, n. 10, p. 1976-1982, 2010.

PATAKY, T. C. One-dimensional statistical parametric mapping in Python. **Computer Methods in Biomechanics and Biomedical Engineering**, 15, n. 3, p. 295-301, 2012.

PATAKY, T. C.; NAOUMA, H.; DONNELLY, C. J. mwarp1d: Manual one-dimensional data warping in python and pyqt. **Journal of Open Source Software**, 4, n. 44, p. 1870, 2019.

PATAKY, T. C.; ROBINSON, M. A.; VANRENTERGHEM, J. Vector field statistical analysis of kinematic and force trajectories. **Journal of Biomechanics**, 46, n. 14, p. 2394-2401, Sep 2013.

PATAKY, T. C.; VANRENTERGHEM, J.; ROBINSON, M. A. Zero- vs. one-dimensional, parametric vs. non-parametric, and confidence interval vs. hypothesis testing procedures in one-dimensional biomechanical trajectory analysis. **Journal of Biomechanics**, 48, n. 7, p. 1277-1285, May 2015.

PENNY, W. D.; FRISTON, K. J.; ASHBURNER, J. T.; KIEBEL, S. J. *et al.* **Statistical parametric mapping: the analysis of functional brain images**. Elsevier, 2011. 0080466508.

PERRY, J.; DAVIDS, J. R. Gait analysis: normal and pathological function. **Journal of Pediatric Orthopaedics**, 12, n. 6, p. 815, 1992.

PRINCE, F.; CORRIVEAU, H.; HÉBERT, R.; WINTER, D. A. Gait in the elderly. **Gait & Posture**, 5, n. 2, p. 128-135, 1997/04/01/ 1997.

SAHA, B.; SRIVASTAVA, D., 2014, **Data quality: The other face of big data**. IEEE. 1294-1297.

SLOOT, L. H.; VAN DER KROGT, M. M. Interpreting Joint Moments and Powers in Gait. *In*: **Handbook of Human Motion**. Cham: Springer International Publishing, 2018. p. 625-643.

SOTELO, M.; EICHELBERGER, P.; FURRER, M.; BAUR, H. *et al.* Walking with an induced unilateral knee extension restriction affects lower but not upper body biomechanics in healthy adults. **Gait & Posture**, 65, p. 182-189, 2018/09/01/ 2018.

TODA, H.; NAGANO, A.; LUO, Z. Age and gender differences in the control of vertical ground reaction force by the hip, knee and ankle joints. **Journal of physical therapy science**, 27, n. 6, p. 1833-1838, 2015.

VAUGHAN, C. L.; DAVIS, B.; O'CONNOR, J. C., 1992, **Dynamics of human gait**.

VAUGHAN, C. L.; DAVIS, B. L.; O'CONNOR, J. C. **Dynamics of human gait**. Human Kinetics, 1992. 0873223705.

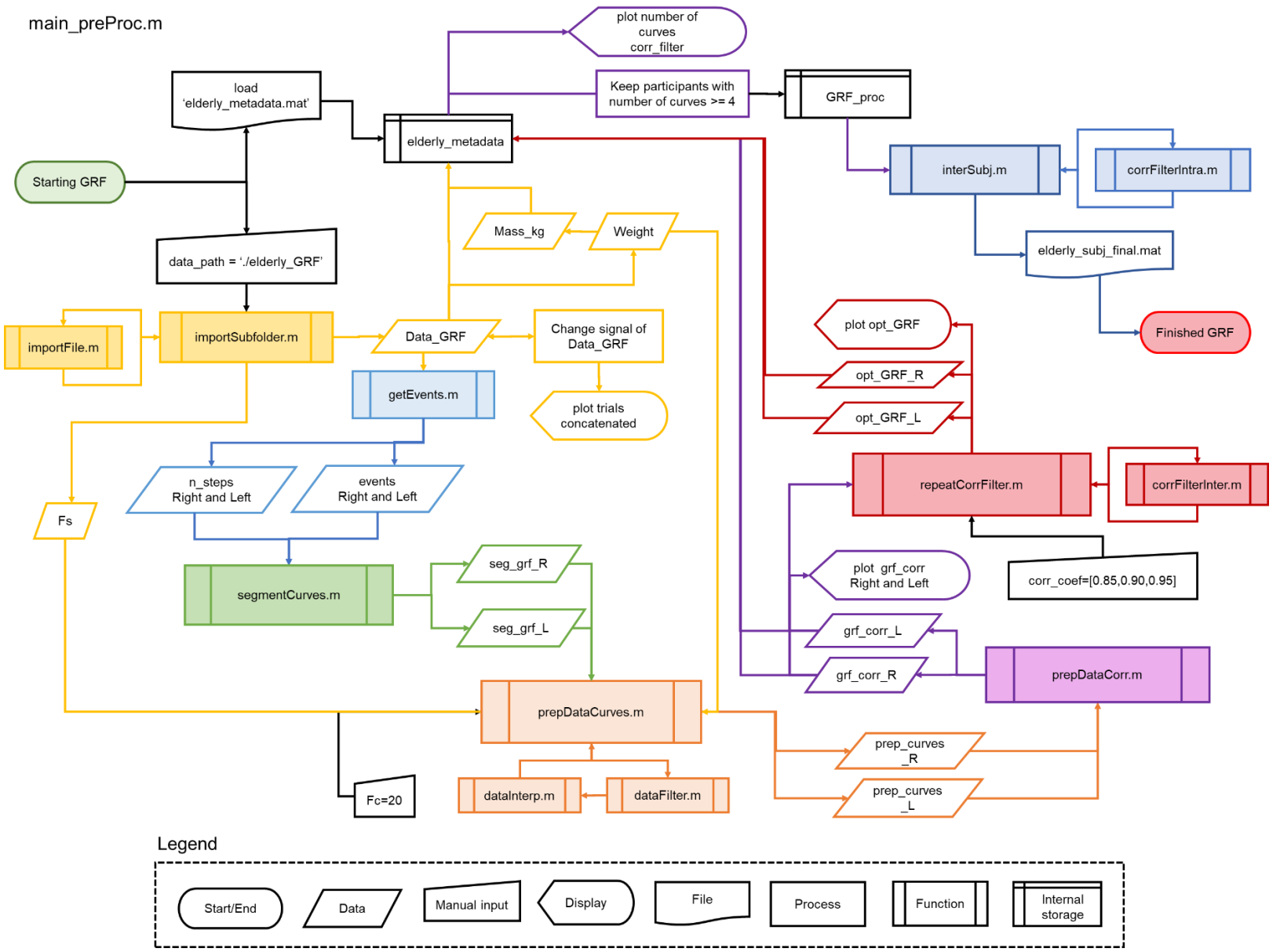
VIEIRA, M. F.; DE BRITO, A. A.; LEHNEN, G. C.; RODRIGUES, F. B. Center of pressure and center of mass behavior during gait initiation on inclined surfaces: A statistical parametric mapping analysis. **Journal of Biomechanics**, 56, p. 10-18, May 2017.

VIEIRA, M. F.; LEHNEN, G. C.; NOLL, M.; RODRIGUES, F. B. *et al.* Use of a backpack alters gait initiation of high school students. **Journal of Electromyography and Kinesiology**, 28, p. 82-89, 2016/06/01/ 2016.

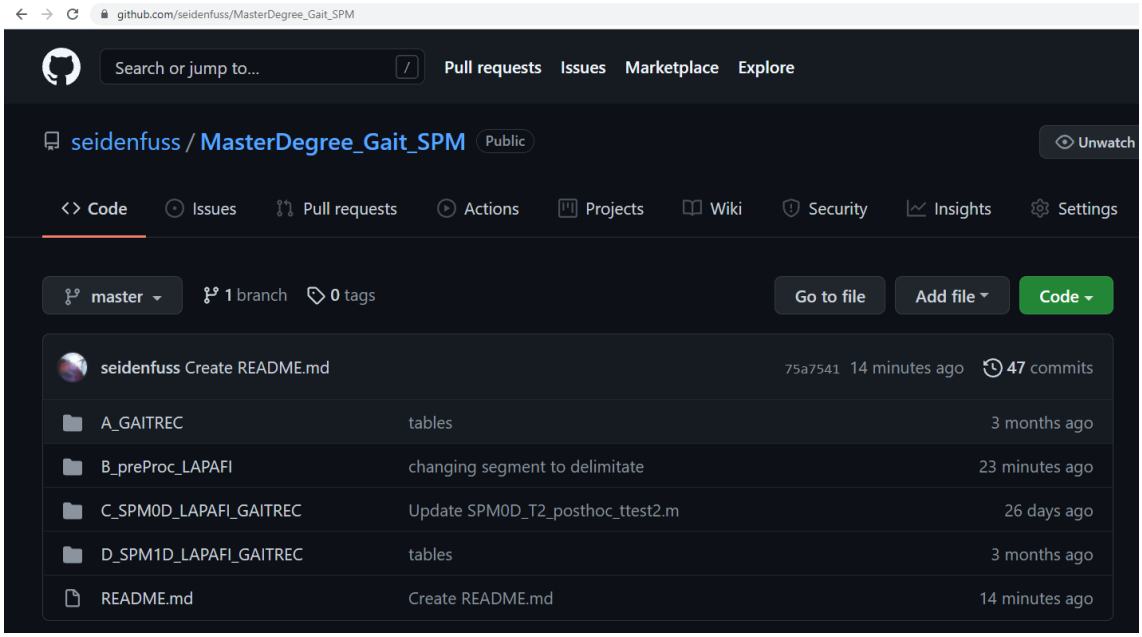
VIEIRA, M. F.; SACCO, I. D. N.; NORA, F.; ROSENBAUM, D. *et al.* Footwear and Foam Surface Alter Gait Initiation of Typical Subjects. **Plos One**, 10, n. 8, Aug 2015.

# 8. APPENDICES

## 8.1. Appendix A - Extended Main Script Flow Diagram for Data Preparation of GRF Data (The Author, 2021).



## 8.2. Appendix B - Data Processing MATLAB Main Scripts and Functions

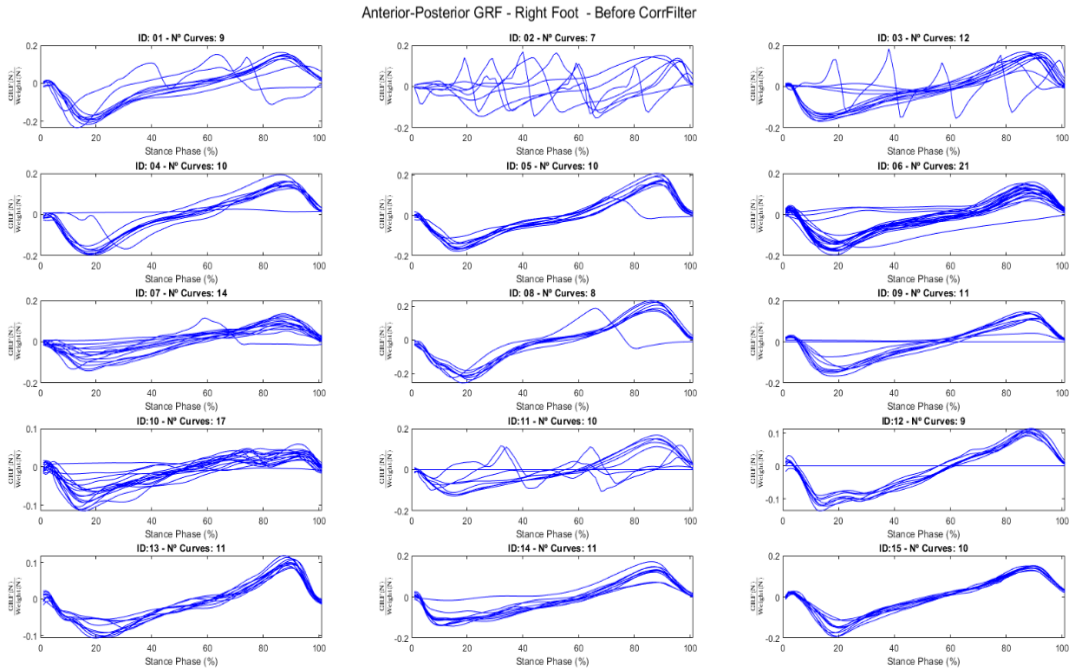



The codes developed for this research and the data processed are available on GitHub: https://github.com/seidenfuss/MasterDegree_Gait_SPM for public access. Any reader is also welcomed to contact me for further information.

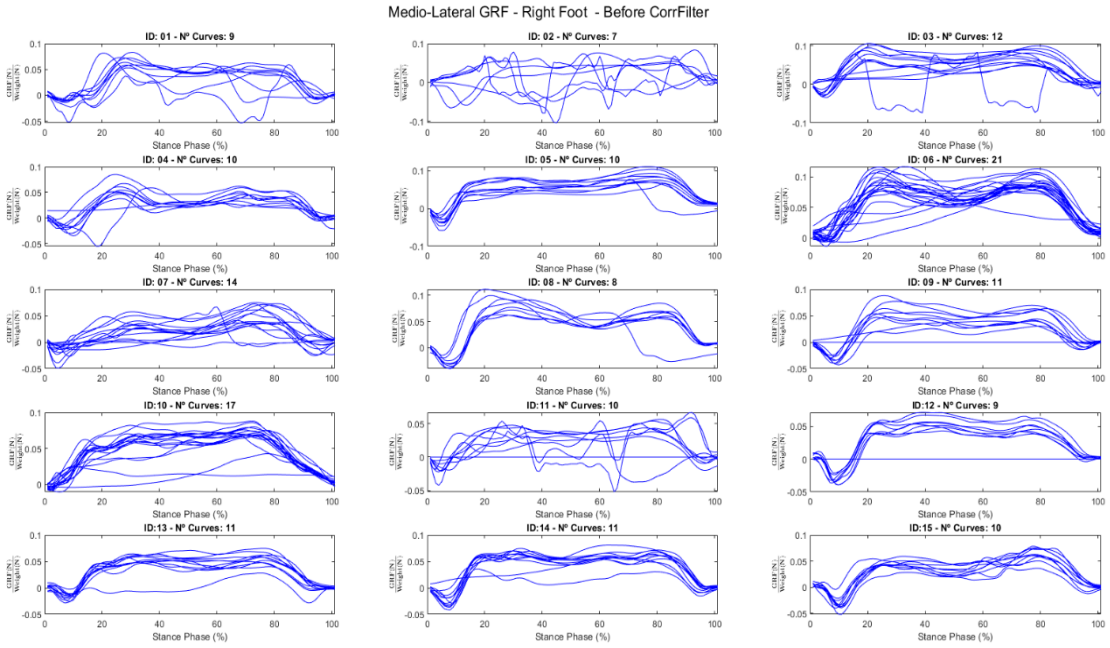
## 8.3. Appendix C - Discrete Values Extracted from Mean GRF Curves for Right and Left Foot

| Variable name | Description |
|---|---|
| min_AP_RL | Minimum value at the anterior posterior GRF curve. |
| locmin_AP_RL | Vector index of the minimum value on anterior posterior GRF curve. |
| max_AP_RL | Maximum value at the anterior posterior GRF curve. |
| locmax_AP_RL | Vector index of the maximum value in anterior posterior GRF curve. |
| inter1_AP_RL | Vector index of lower limit for integral calculation. |
| mid_AP_RL | Vector index of the middle point where the anterior posterior curve crosses the zero, transitioning from breaking to acceleration phases of stance. |
| Inter2_AP_RL | Vector index of superior limit for integral calculation. |
| peak1_V_RL | First peak of vertical GRF: maximum value between the begin and half of the curve. |
| locpeak1_V_RL | Vector index of the first peak of vertical GRF |
| peak2_V_RL | Second peak of vertical GRF: maximum value between the half of the curve and its end. |
| locpeak2_V_RL | Vector index of the second peak of vertical GRF. |
| valley_V_RL | Vertical GRF valley: between the first and second GRF peaks vector indices. |
| locvalley_V_RL | Vector index of the vertical GRF valley. |
| valley1_ML_RL | Valley of medial-lateral GRF curve: the minimum value between the begin and half of the curve. |
| valley1_ML_RL | Vector index of the medial-lateral GRF valley. |
| peak1_ML_RL | First peak of medial-lateral GRF curve: maximum value between the begin and half of the curve. |
| locpeak1_ML_RL | Vector index of first peak of medial-lateral GRF curve. |
| peak2_ML_RL | Second peak of medial-lateral GRF curve: maximum value between half of the curve and its end. |
| locpeak2_ML_RL | Vector index of second peak of medial-lateral GRF curve. |
| inter1_ML_RL | Lower limit for integral calculation of medial-lateral GRF curve. |
| inter2_ML_RL | Superior limit for integral calculation of medial-lateral GRF curve. |
| total_imp_AP_RL | Numerical integration using trapezoidal method between inter1_AP_RL and inter2_AP_RL |
| breaking_imp_AP_RL | Numerical integration using trapezoidal method between inter1_AP_RL and mid_AP_RL |
| propulsion_imp_AP_RL | Numerical integration using trapezoidal method between mid_AP_RL and inter2_AP_RL |
| total_imp_V_RL | Numerical integration using trapezoidal method between first and last vector indeces (1 and 101). |
| firsthalf_imp_V_RL | Numerical integration using trapezoidal method between first point (vector index=1) and locvalley_V_RL. |
| secondhalf_imp_V_RL | Numerical integration using trapezoidal method between locvalley_V_RL and last point (vector index = 101). |
| earlystance_imp_V_RL | Numerical integration using trapezoidal method between first point (vector index=1) and locpeak1_V_RL. |
| midstance_imp_V_RL' | Numerical integration using trapezoidal method between locpeak1_V_RL and locpeak2_V_RL. |
| latestance_imp_V_RL' | Numerical integration using trapezoidal method between locpeak2_V_RL and the last point (vector index = 101). |
| total_imp_ML_RL | Numerical integration using trapezoidal method between inter1_ML_RL and inter2_ML_RL |

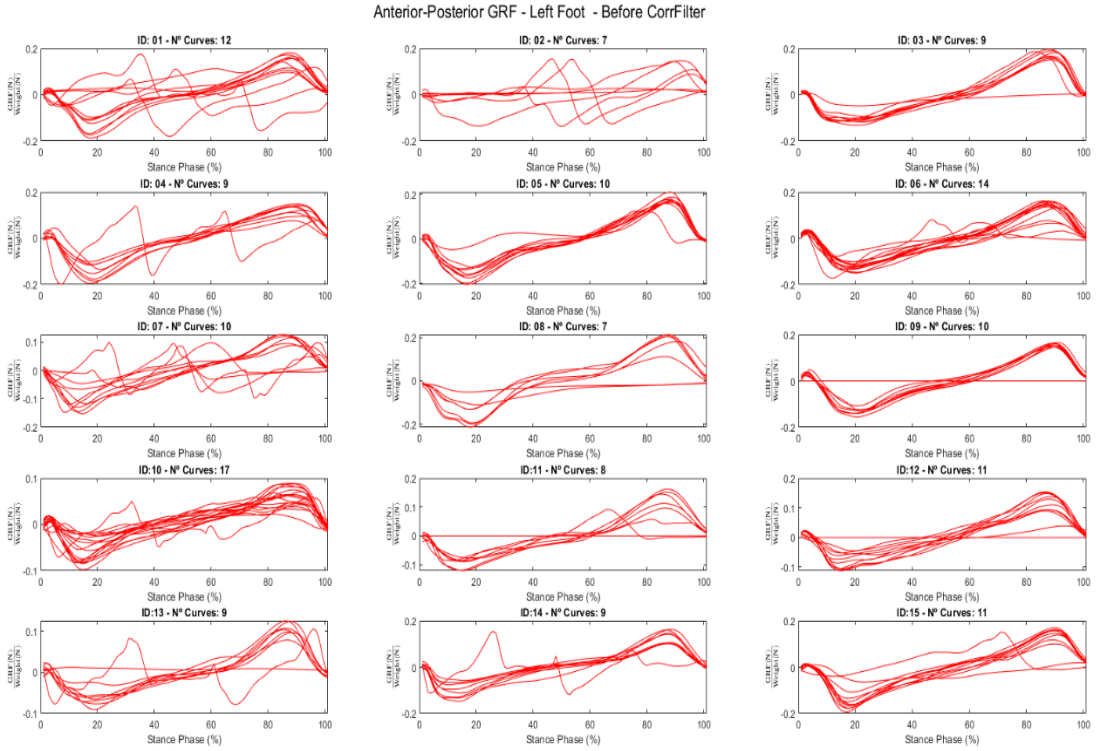
## 8.4. **Appendix D - Prepared Curves (shear GRF vector components): delimited, noise removed, down sampled, and interpolated curves**



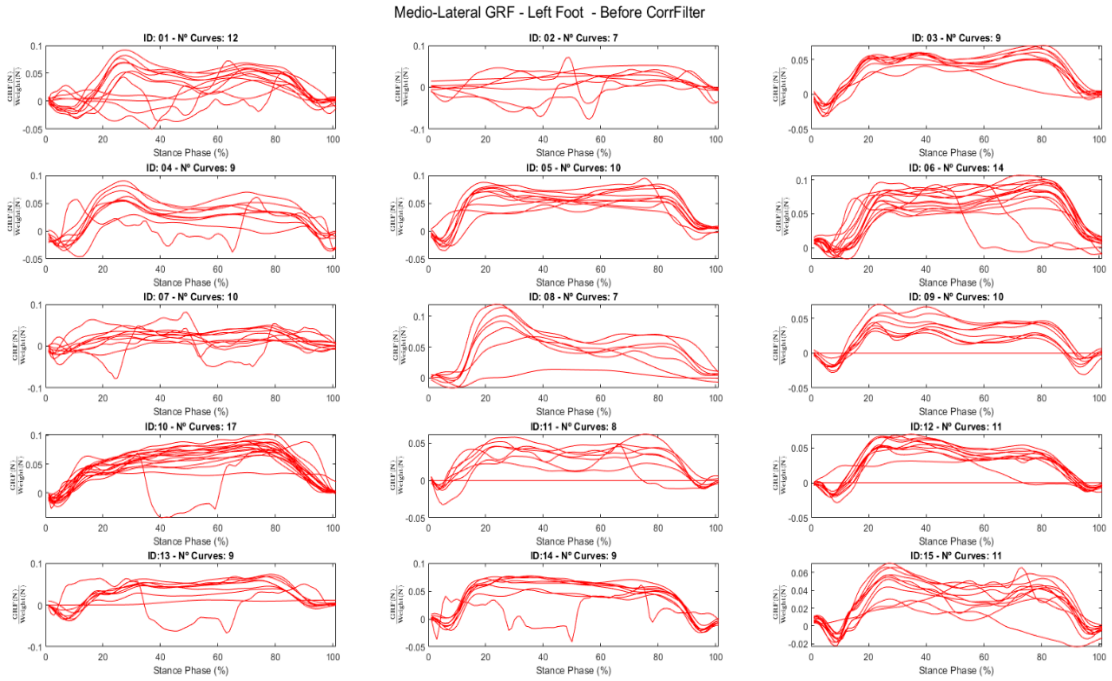

*Anterior posterior GRF component of right foot (The Author, 2021).*




*Medial-lateral GRF component of right foot (The Author, 2021).*

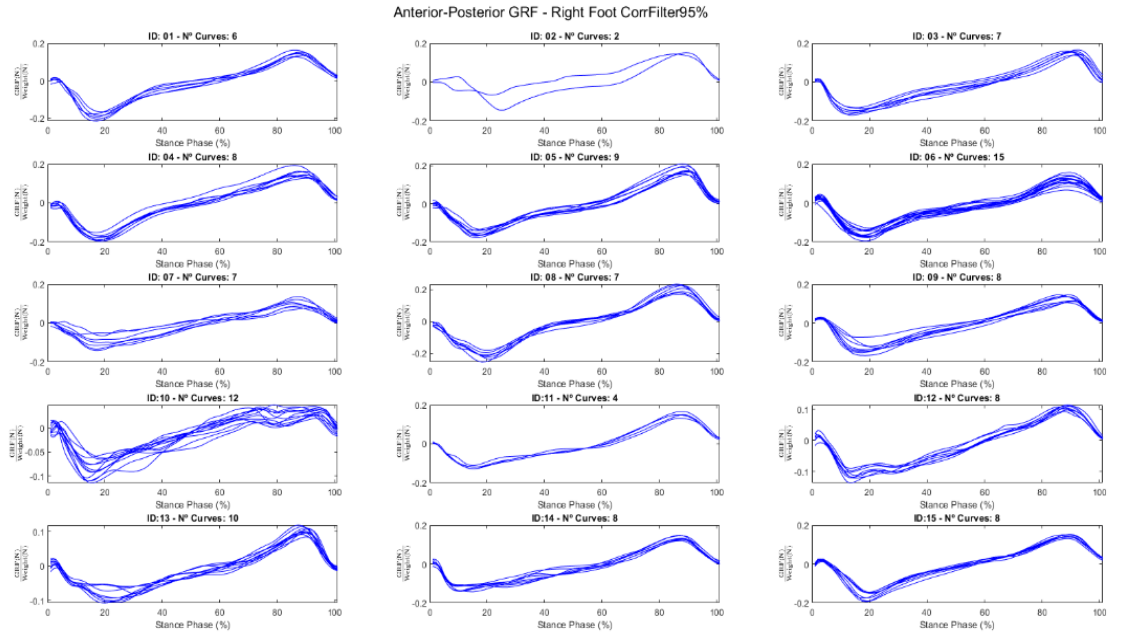



*Anterior posterior GRF component of left foot. (The Author, 2021).*

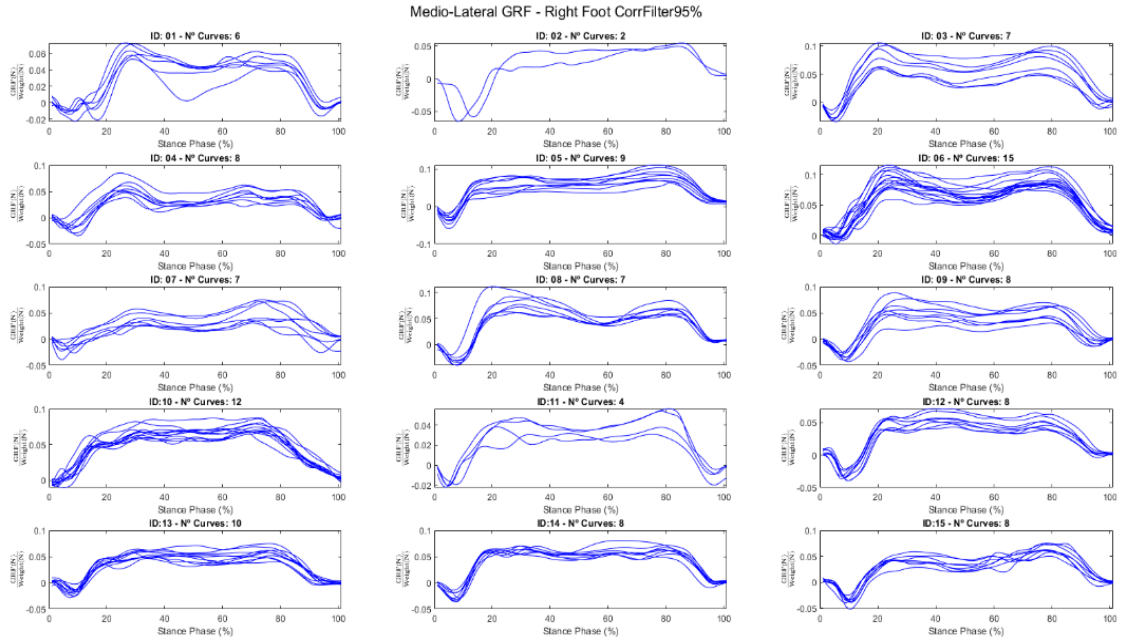



*Medial-lateral GRF component of left foot. (The Author, 2021).*

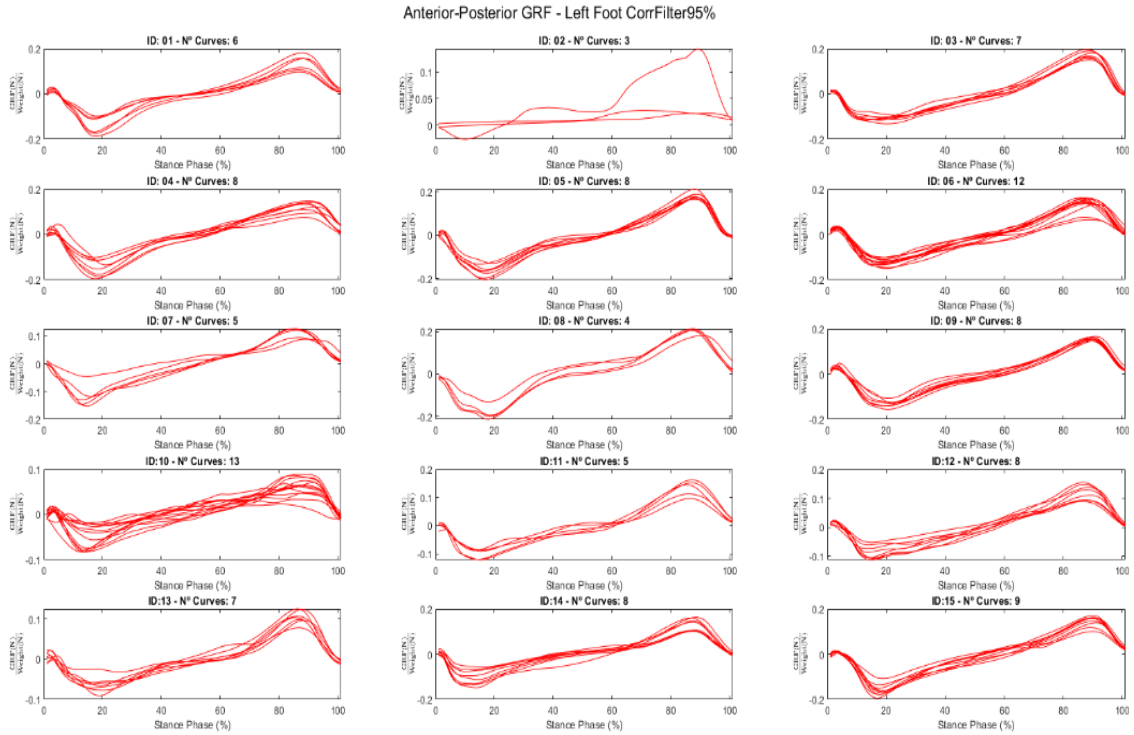
## 8.5. Appendix E – Intrasubject Filter Results: Shear GRF Components after Correlation Filter (threshold of 95%)



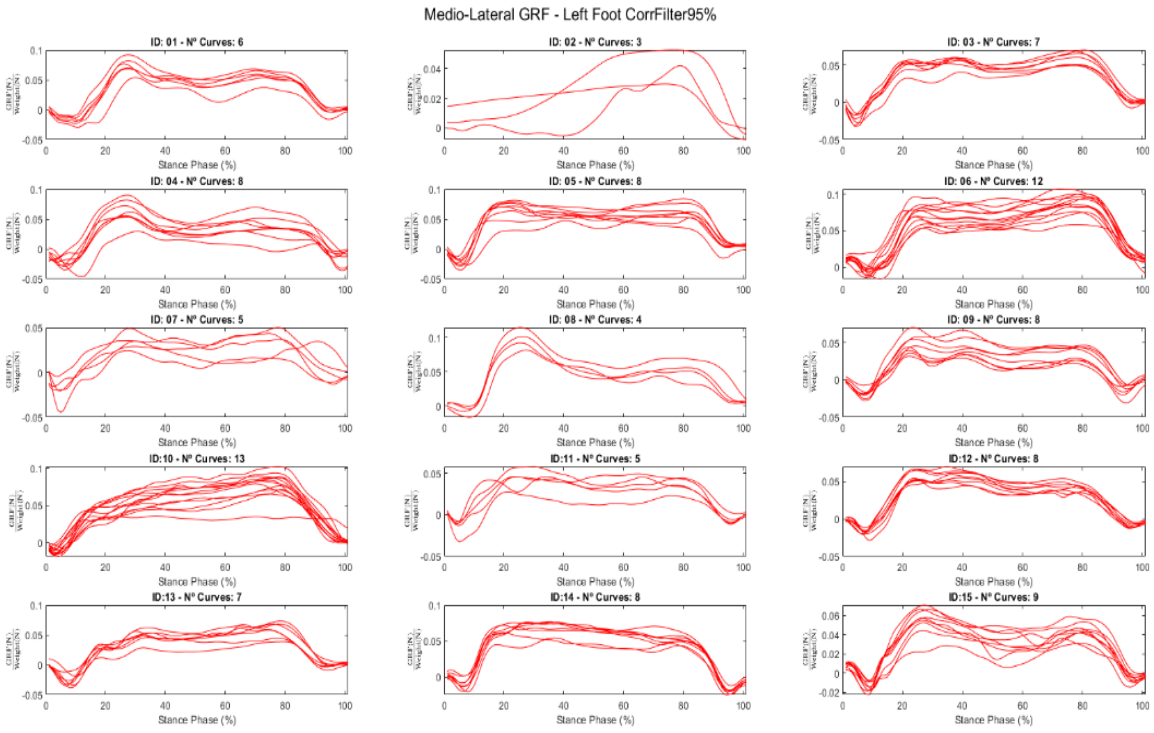

*Anterior posterior GRF component right foot (The Author, 2021).*




*Medial-Lateral GRF component right foot (The Author, 2021).*

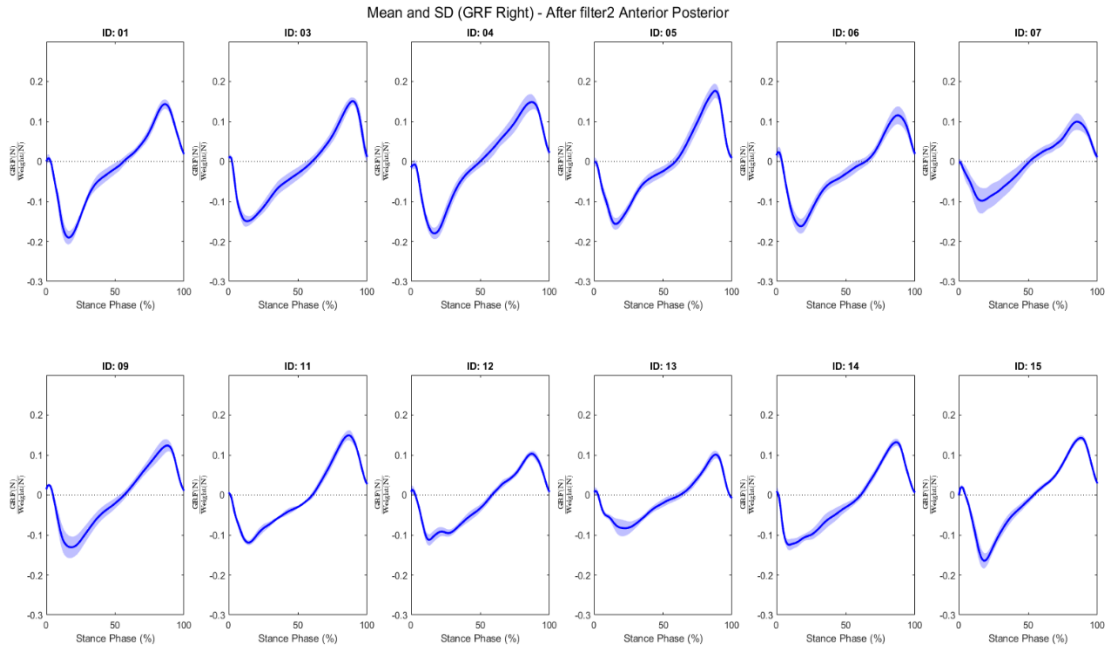



*Anterior posterior GRF left foot (The Author, 2021).*

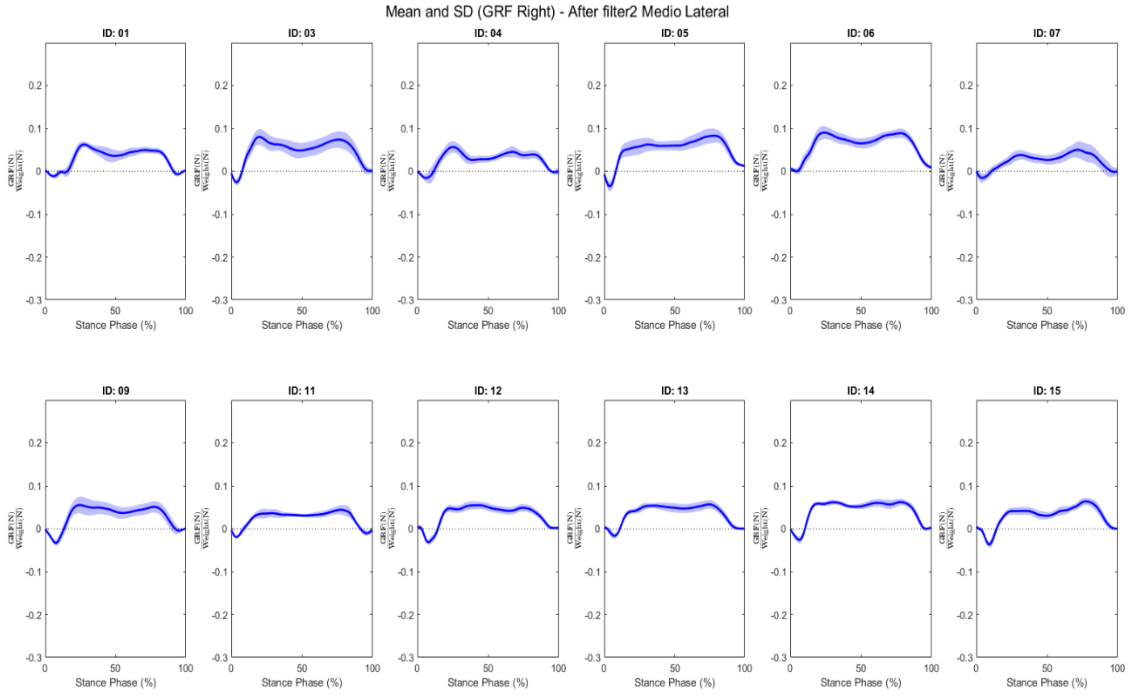



*Medial-Lateral GRF component left foot (The Author, 2021).*

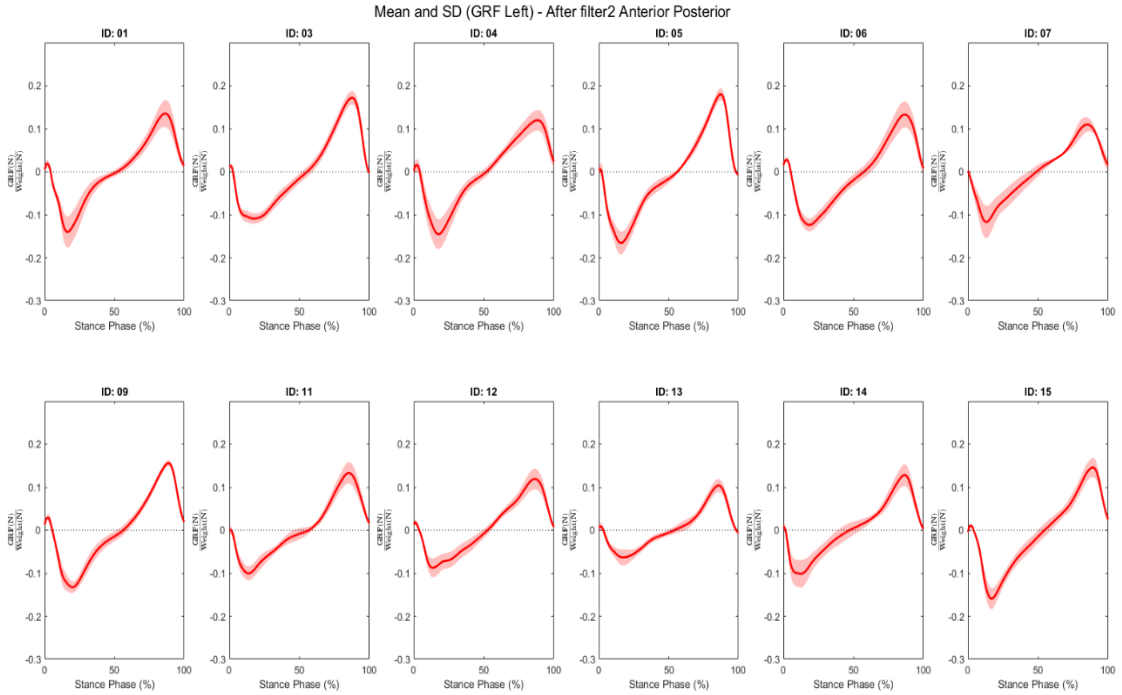
## 8.6. Appendix F - Intersubjects Filter Results: Shear GRF Components



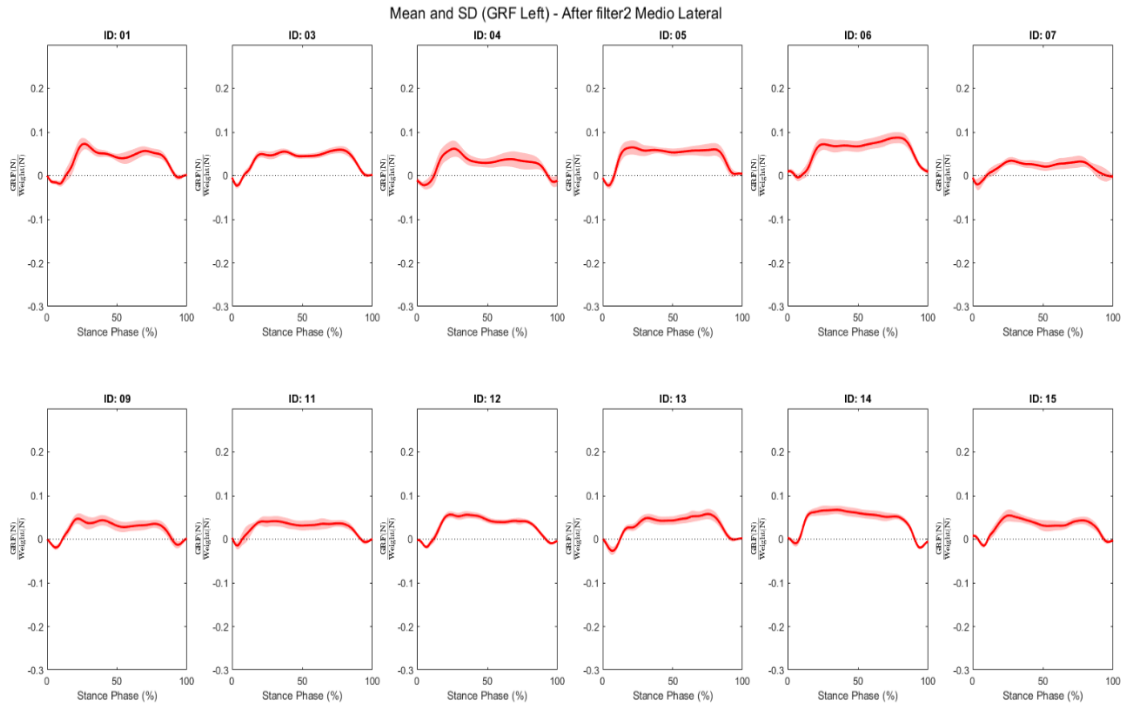

*Anterior posterior GRF mean curves right foot (The Author, 2021).*




*Medial-lateral GRF mean curves right foot (The Author, 2021).*

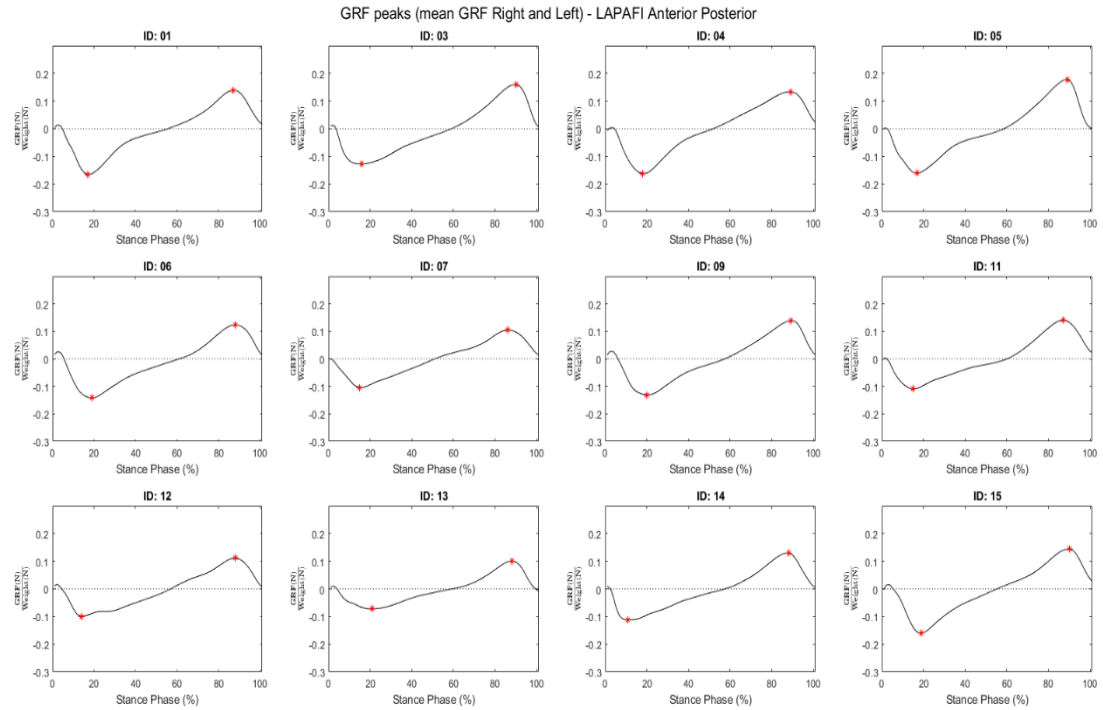



*Anterior posterior GRF mean curves left foot (The Author, 2021).*

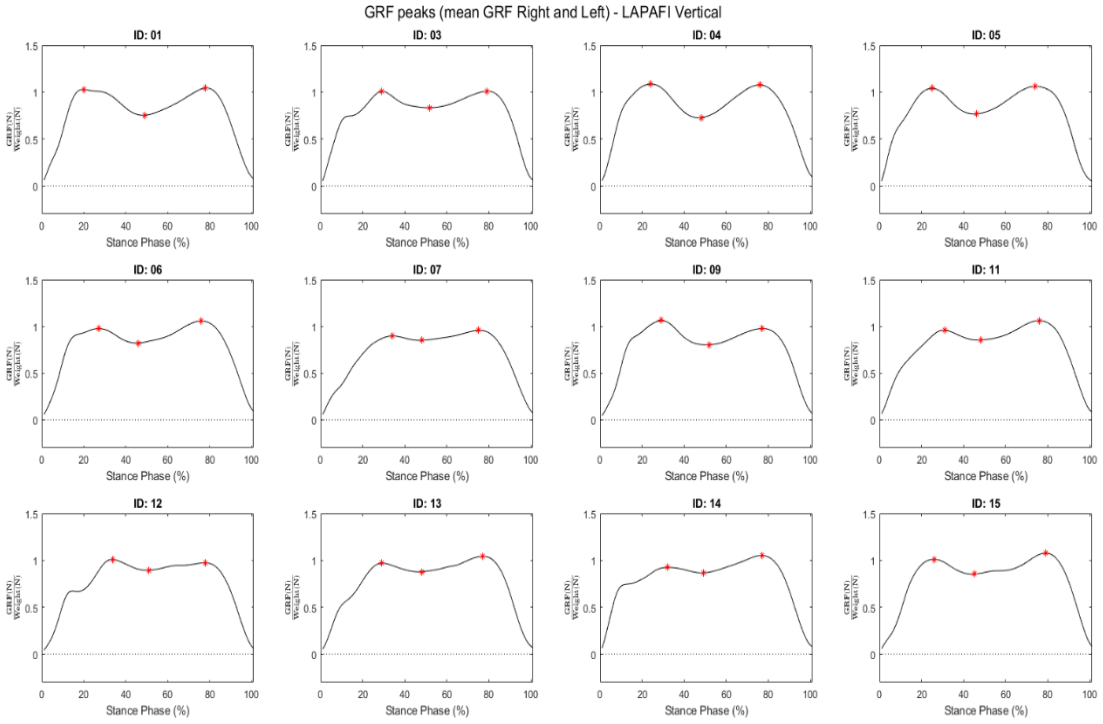



*Medial-lateral GRF mean curves left foot (The Author, 2021).*

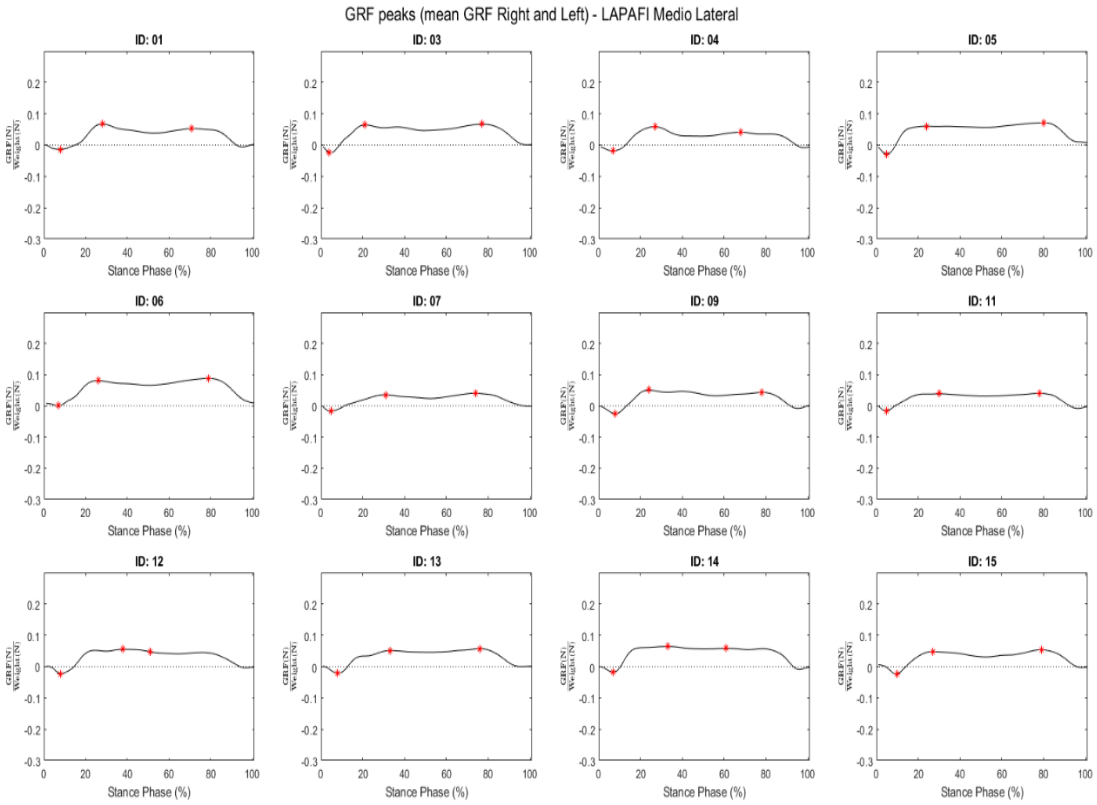
## 8.7. Appendix G - GRF Peaks and Valleys - LAPAFI




*Peaks GRF anterior posterior LAPAFI (The Author, 2021).*

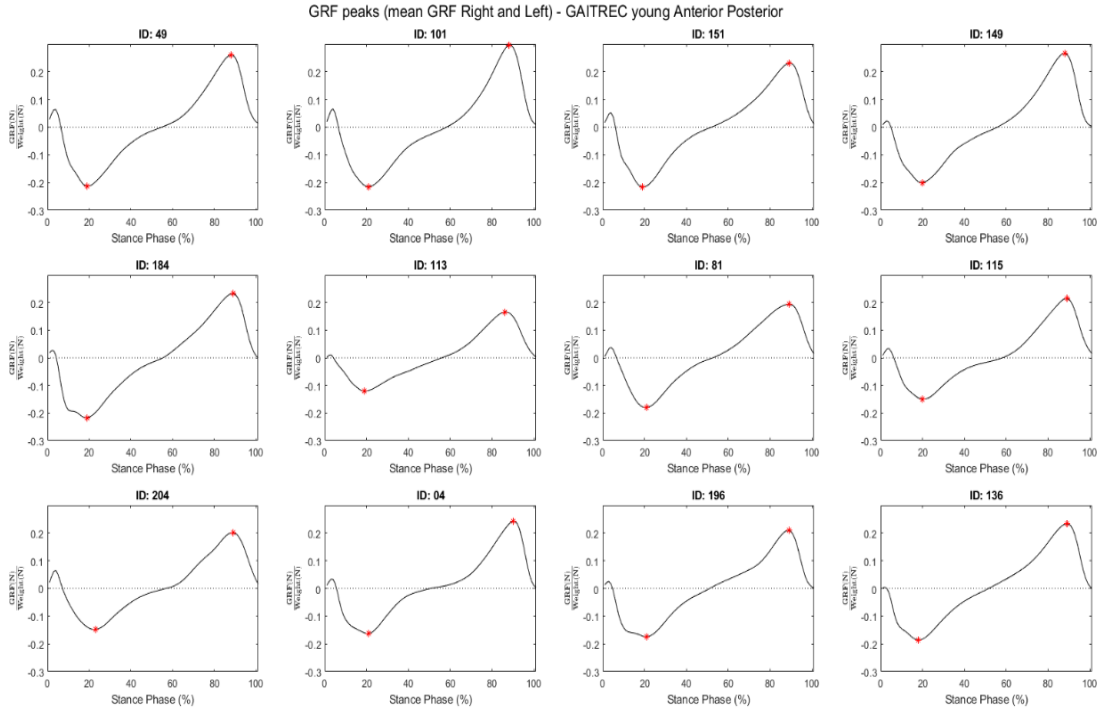



*Peaks GRF vertical LAPAFI (The Author, 2021).*

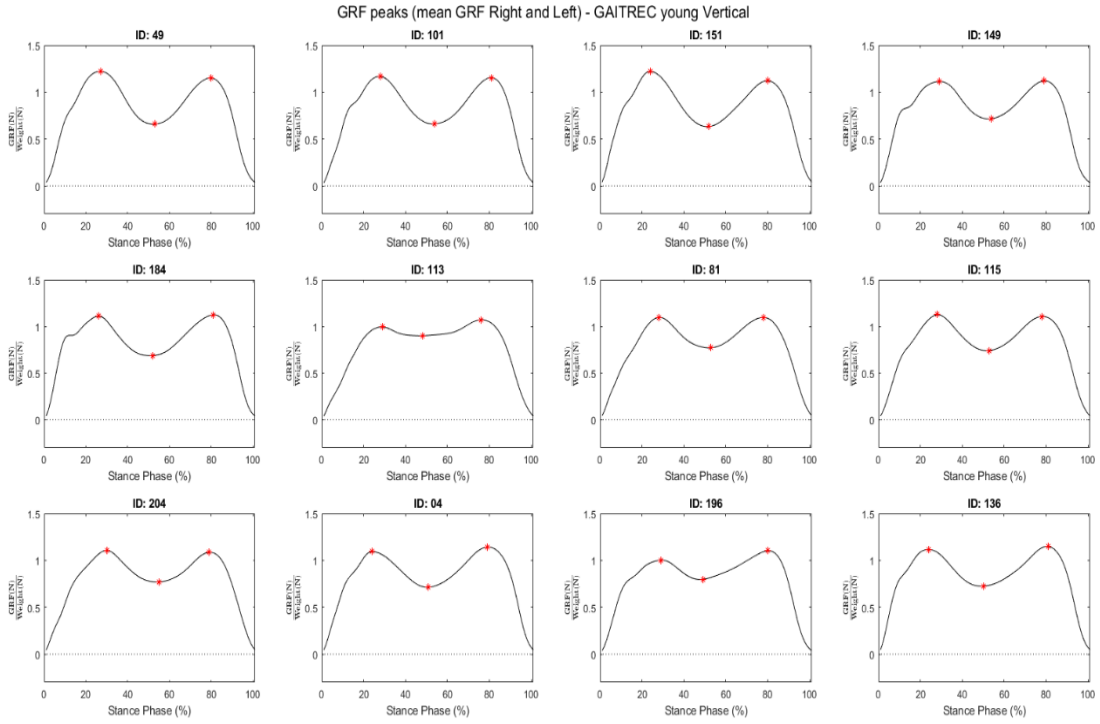



*Peaks GRF medial-lateral LAPAFI (The Author, 2021).*

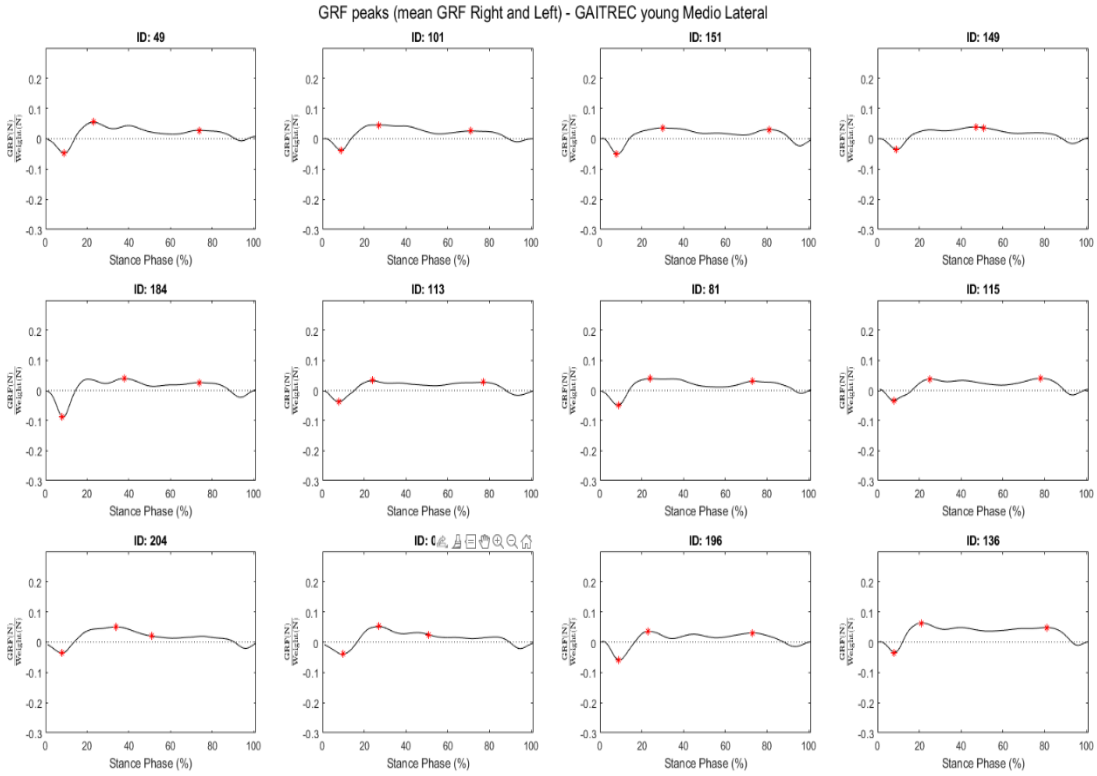
## 8.8. Appendix H - GRF Peaks and Valleys - GAITREC



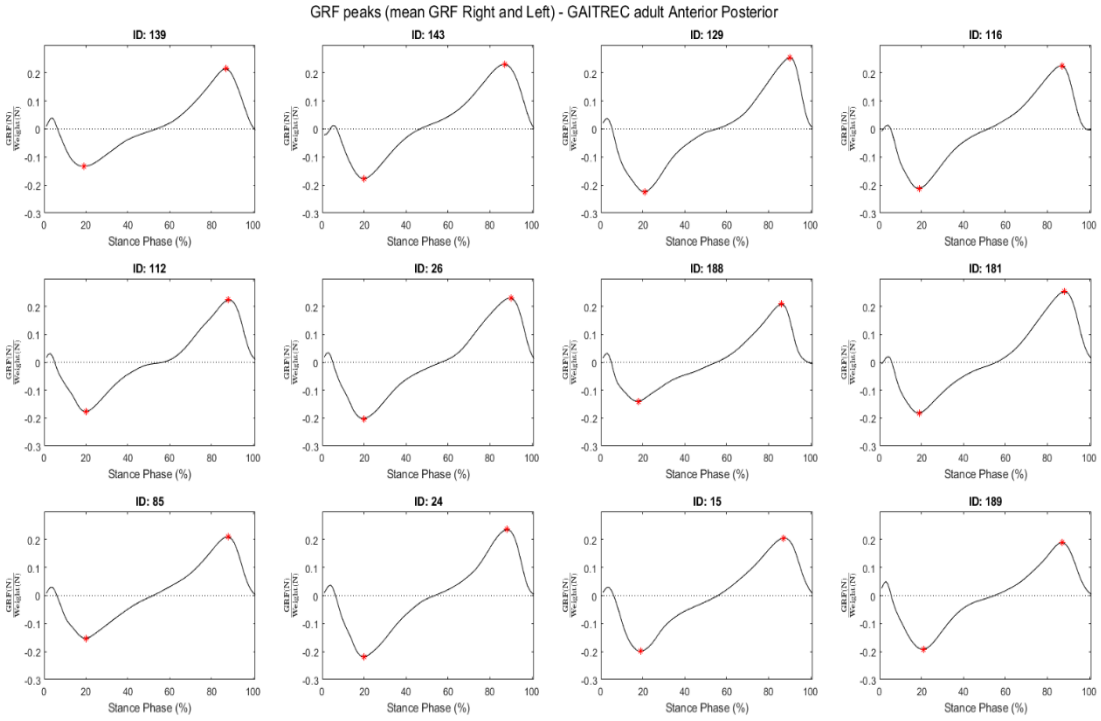

*Peaks GRF - GaitRec young - Anterior posterior. (The Author, 2021).*



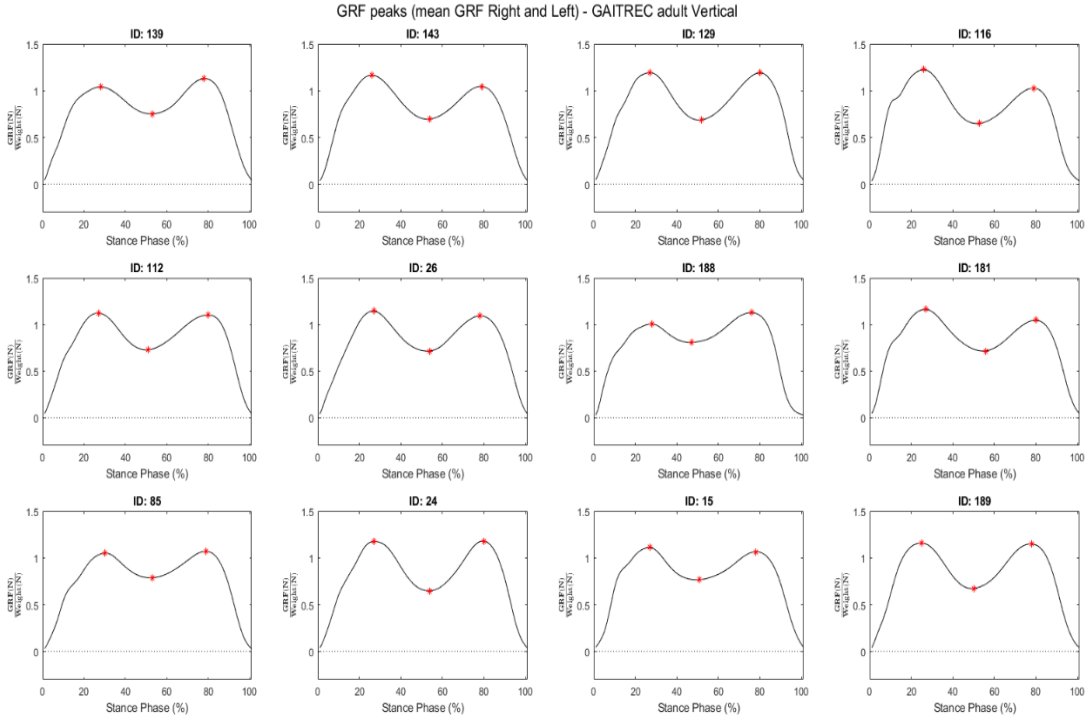

*Peaks GRF - GaitRec young – Vertical. (The Author, 2021).*



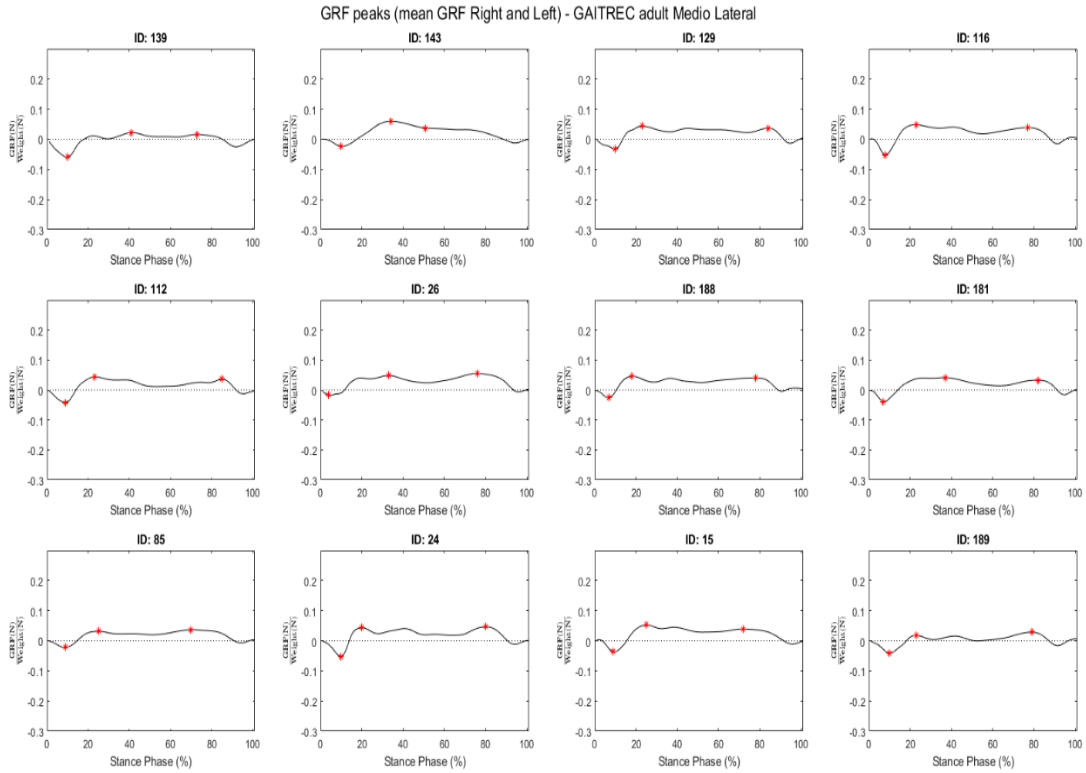

*Peaks GRF - GaitRec young - Medial-Lateral. (The Author, 2021).*



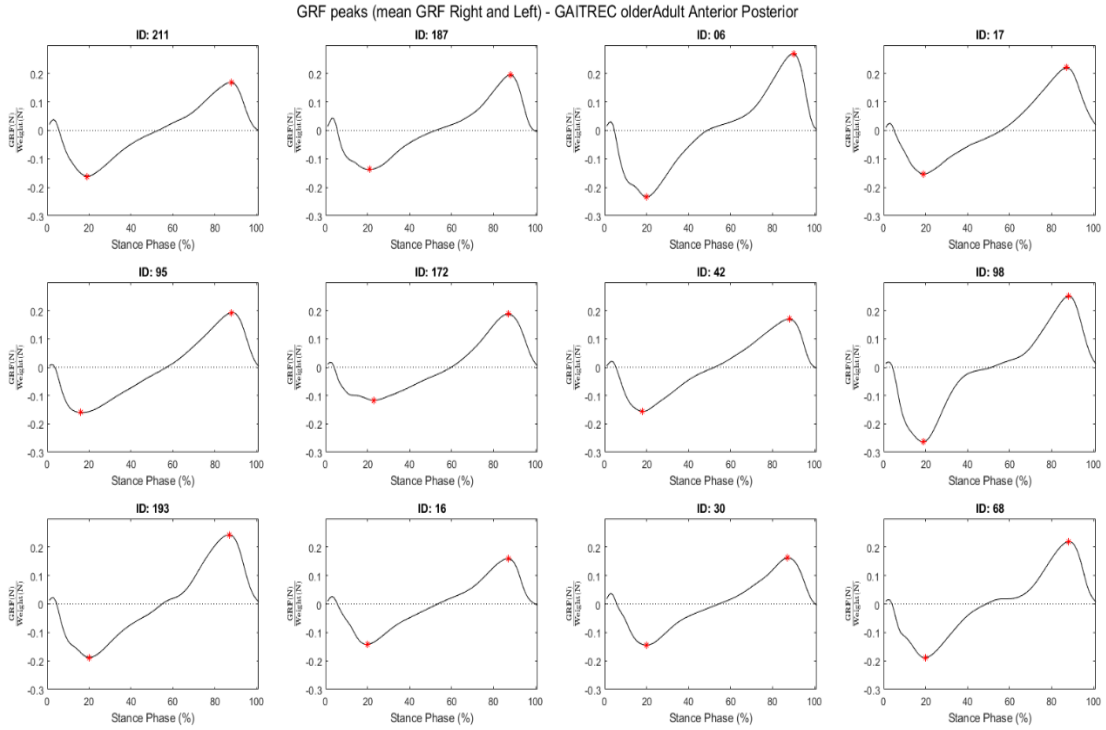

*Peaks GRF - GaitRec adult - Anterior posterior. (The Author, 2021).*



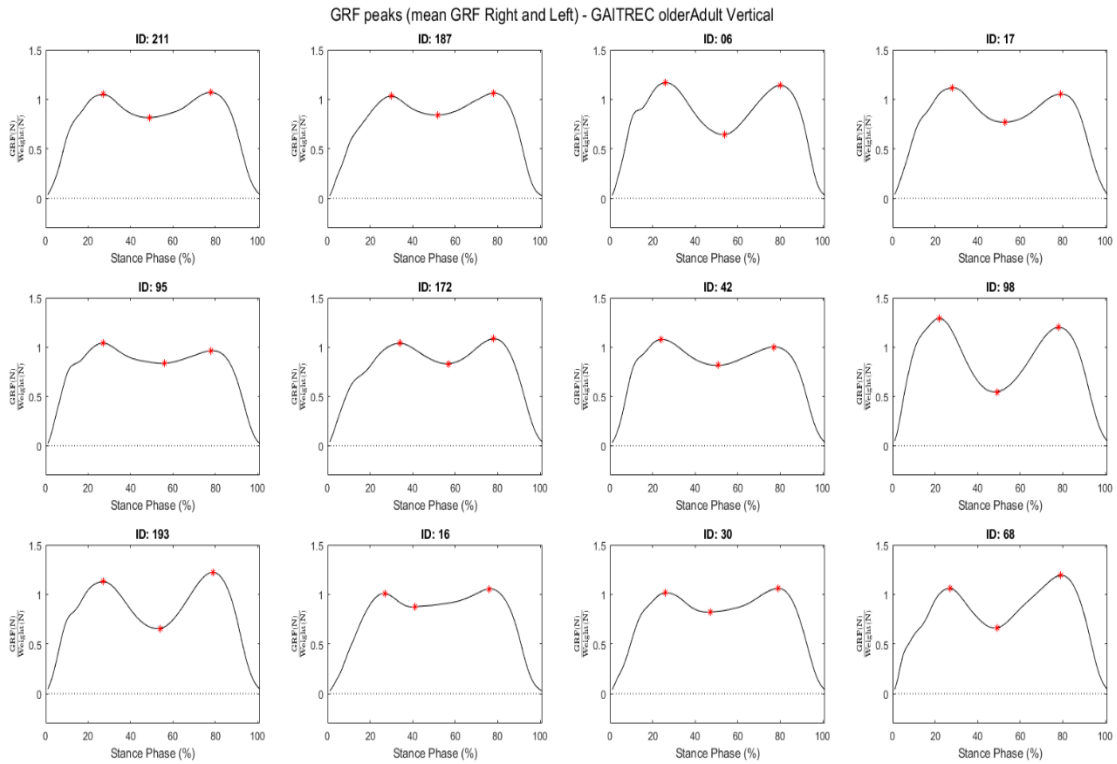

*Peaks GRF - GaitRec adult – Vertical. (The Author, 2021).*



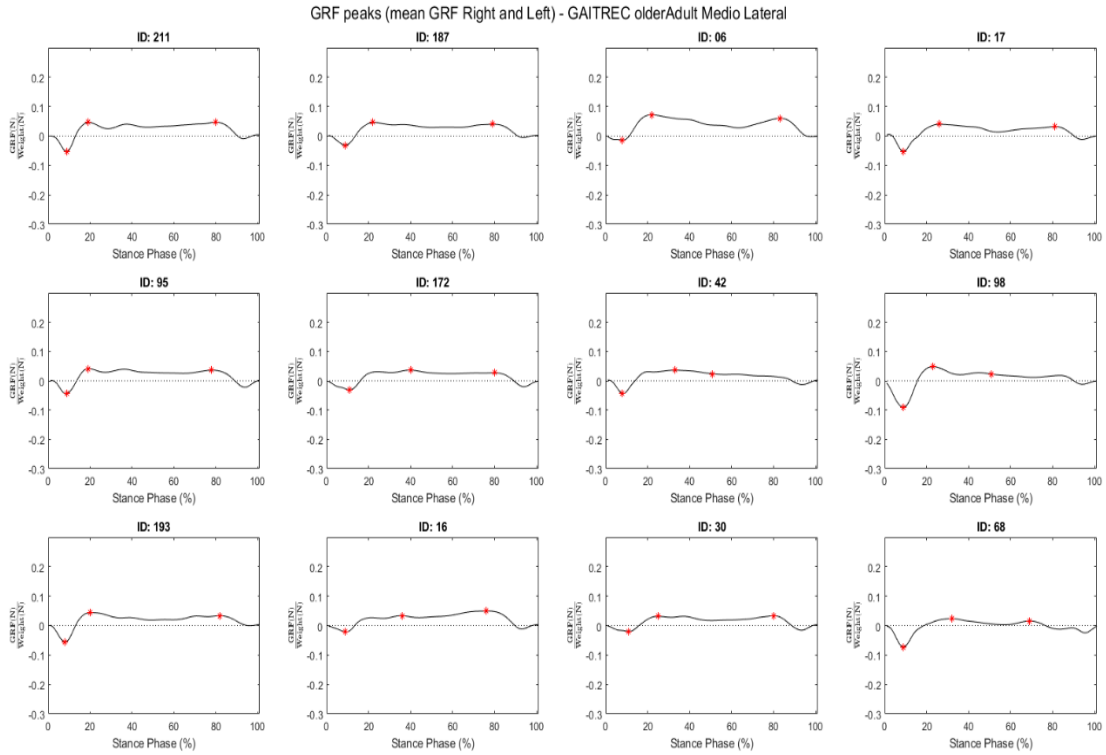

*Peaks GRF - GaitRec adult - Medial-lateral. (The Author, 2021).*




*Peaks GRF - GaitRec older adult - Anterior posterior. (The Author, 2021).*




*Peaks GRF GaitRec older adult Vertical. (The Author, 2021).*

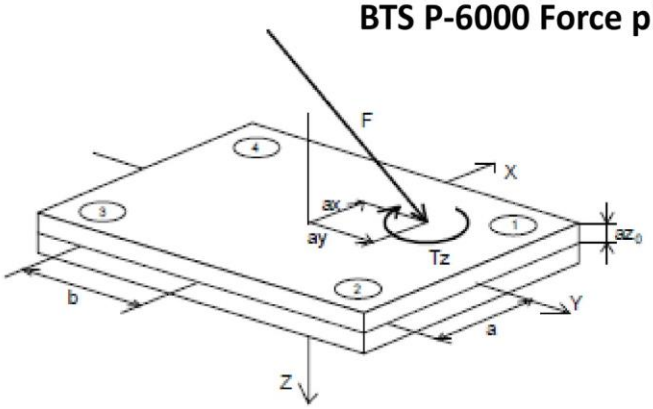



*Peaks GRF GaitRec older adult - Medial-lateral. (The Author, 2021).*

# 9. ANNEXES

## 9.1. Annex A - BTS Bioengineering force plate formulae (Model P-6000)

### BTS P-6000 Force plate Formulae




a = 164 mm
b = 264 mm
z0 = 26.5 mm

Force plate Output Signals

| Output signal | Channel | Description |
|---|---|---|
| fx12 | 1 | Force in X-direction measured by sensor 1 + sensor 2 |
| fx34 | 2 | Force in X-direction measured by sensor 3 + sensor 4 |
| fy14 | 3 | Force in Y-direction measured by sensor 1 + sensor 4 |
| fy23 | 4 | Force in Y-direction measured by sensor 2 + sensor 3 |
| fz1 ... fz4 | 5 ... 8 | Force in Z direction measured by sensor 1 ... 4 |

| Parameter | Calculation | Description |
|---|---|---|
| Fx | = fx12 + fx34 | Medio-lateral force 1) |
| Fy | = fy14 + fy23 | Anterior-posterior force 1) |
| Fz | = fz1 + fz2 + fz3 + fz4 | Vertical force |
| Mx | = b * (fz1 + fz2 - fz3 - fz4) | Plate moment about X-axis 3) |
| My | = a * (-fz1 + fz2 + fz3 - fz4) | Plate moment about Y-axis 3) |
| Mz | = b * (-fx12 + fx34) + a * (fy14 - fy23) | Plate moment about Z-axis 3) |
| Mx' | = Mx + Fy*az0 | Plate moment about top plate surface 2) |
| My' | = My - Fx*az0 | Plate moment about top plate surface 2) |
| ax | = -My' / Fz | X-Coordinate of force application point (COP) 2) |
| ay | = Mx' / Fz | Y-Coordinate of force application point (COP) 2) |
| Tz | = Mz - Fy * ax + Fx * ay | Free moment, Vertical torque, „Frictional" torque |

1) Walking direction is positive Y-axis

2) az0 = top plane offset (negative value)

3) a, b = sensor offset (positive values)

## 9.2. Annex B - Main features of each force platform module

| | |
|---|---|
| Dimension Single Module<br>(equivalent to 1 traditional platform) | Sensitive area 60 x 40cm<br>minimum height 6cm |
| Interface | LAN (10/100 Ethernet) |
| Signal Output | Digital |
| Power Supply | PoE with proprietary switch |
| Capacity (X and Y) for each module | Up to ±8000 N |
| Capacity (Z) for each module | Up to 8000 N |
| Sensitivity/Resolution | 16 bit over selected range |
| Sensitivity deviation over plate surface | <1,0% Full Scale Output |
| Hysteresis | <0,2% Full Scale Output |
| Linearity | <0,2% Full Scale Output |
| Sensing elements | Patented strain gauge architecture |
| Amplifier | Built-in |
| Mounting Hardware | Not required |
| Protection degree | IP42 |
| Compliance to Standards | Safety: EN 60601-1<br>EMC: EN 60601-1-2 |

## 9.3. Annex C - GaitRec dataset experimental protocol aspects

| Aspect | Description |
|---|---|
| Data collection | Bilateral GRF at self-selected speed recorded between 2007 and 2018. |
| Force platforms | Two centrally embedded force plates (Kistler, Type 9281B12, Winterthur, CH) on a 10 m walkway covered with the same material. |
| Valid records | Subjects walked until (usually) ten valid recordings were available - valid recordings were defined by the assessor when the participant walked naturally and there was a clean foot strike on each force plate. |
| Speed | Healthy controls walked at three different walking speeds (mean and standard deviation, m/s): slow 0.98 (0.14), self-selected 1.27 (0.13), and fast 1.55 (0.15). |
| Shod condition | Healthy controls walked either barefoot or with their usual shoes. |
| Health condition | Participants underwent rigorous physical examination by a physician prior to the gait analysis session. |
| Digitalization of signal | Conversion from analog to digital signals: sampling rate of 2000 Hz and a 12-bit analog-digital converter (DT3010, Data Translation Incorporation, Marlboro, MA, USA) with a signal input range of ±10 V. |
| Coordinate system | Local force plate coordinate system (reaction-orientated). Orientation of the medio-lateral and anterior-posterior signals were uniformed: medial and anterior forces are always represented as positive values. |
| Raw data | Raw data: Signals are available down-sampled to 250 Hz. Threshold of 25 N (to avoid noise and signal peaks at the beginning and end of signals). |
| Processed data | Processed data: ready to use data. |
| Digital filter | Filtering: 2nd order low-pass butterworth filter with a cut-off frequency of 20 Hz. |
| Normalization | The data was time-normalized to 100% stance (i.e. 101 points). Amplitude values of the force components were expressed as a multiple of body weight (BW). |
| Software | All processing steps were performed in MATLAB 2019a (The MathWorks Inc., Natick, MA, USA). |


Pontifícia Universidade Católica do Rio Grande do Sul
Pró-Reitoria de Graduação
Av. Ipiranga, 6681 - Prédio 1 - 3º. andar
Porto Alegre - RS - Brasil
Fone: (51) 3320-3500 - Fax: (51) 3339-1564
E-mail: prograd@pucrs.br
Site: www.pucrs.br