Test-Retest reliability of kinematic measurements in gait analysis of obese adolescents

Master Thesis

For attainment of the academic degree of
Master of Science in Engineering (MSc)

in the Master Program Digital Healthcare
at St. Pölten University of Applied Sciences

by
MSc Bojan Makivic BA
dh151829

First advisor: FH-Prof. Dr. Brian Horsak

St. Pölten, 17.08.2017

Declaration

I declare that I have developed and written the enclosed Master Thesis completely by myself, and have not used sources or means without declaration in the text. Any thoughts from others or literal quotations are clearly marked. This work was not used in the same or in a similar version to achieve an academic grading or is being published elsewhere.

....

Place, Date Signature

Abstract

Introduction: Clinical gait analysis is considered as gold standard for evaluating gait abnormalities in patients. In the 3D clinical gait analysis (3DGA) practice, decisions are typically made based on all the strides collected while the researcher might be more interested in the patient's most representative stride. Important information from the shape of the curve might be lost if the averaging method is utilized whereas visual inspection of the most representative gait curve is a time-consuming job. Therefore, the purpose of this study was to investigate if it makes a difference in kinematic variables if one uses the approach to identify a most representative trial derived from functional median distance depth or by using an average across five trials curve.

Methods: 10 adolescent participants (2 girls and 8 boys) with an age-based body mass index (BMI) above 97th percentile (mean ± SD: age: 14.6 ± 2.8 years, height: 169.3 ± 11.3 cm, body mass: 99.2 ± 21.7 kg; BMI: 34.2 ± 3.9 kg/m²) were selected and administrated to two 3DGA sessions carried out by the same assessor. The multivariate analysis proposed by Sangeux and Polak was used to obtain a most representative trial (MRT) from five trials performed by each subject. An average curve (AVG) was calculated for each kinematic parameter as well. Waveform similarity between the most representative trial and the average curve was estimated with the linear fit model (LFM) and root mean square deviation (RMSD). Additionally, the test-retest reliability was quantified by calculating the standard error of measurement (SEM).

Results: The LFM results for each plane indicated good waveform similarity between AVG and MRT. For all three planes the average linear relationship (R^2) was above 0.9. Among all other joints, the foot progression angle (FPA) showed the lowest R^2=0,82 for single kinematic parameters while knee and hip angle in sagittal plane showed the highest $R^2=1$. The RMSD for the frontal, sagittal and transversal plane were on average 0.58° ± 0.2°, 1.13 ± 0.56° and 1.28 ± 0.4° for the test and 0.65° ± 0.19°, 1.25° ± 0.72° and 1.34° ± 0.66° for the retest session, respectively. The SEM values between test and retest were below 5° for all kinematic parameters indicating good to moderate reliability.

Conclusion: The results suggest the usefulness of MRT in clinical gait analysis due to good similarity to the averaged curve across multiple trials (AVG) estimated by LFM and RMSD. The possible advantage of MRT lies in its simplicity and retention of relevant data/shape information. Additionally, SEM analysis showed acceptable test-retest reliability.

Kurzfassung

Einleitung: Die klinische Ganganalyse wird in der Literatur als Goldstandart zur Evaluierung von pathologischen Gangbildern bei PatientInnen beschrieben. In der Praxis werden bei 3D Ganganalysen klinische Entscheidungen typischerweise aufgrund aller aufgezeichneten Schritte getroffen, obwohl man möglicherweise mehr Interesse an einem repräsentativen Schritt haben könnte. Es wird vermutet, dass wichtige Informationen aufgrund der Mittelung aller Daten verloren gehen könnten. Eine visuelle Inspektion der einzelnen Schritte könnte dem entgegenwirken, ist in der Praxis aber mit erheblichem Zeitaufwand verbunden und daher nicht praktikabel. Aus diesem Grund hat sich die vorliegende Studie zum Ziel gesetzt, zu vergleichen, ob die gemittelten kinematischen Daten sich von einem auf Basis statistischer Analyseverfahren ausgewählten repräsentativen Schritt unterscheiden.

Methoden: 10 Jugendliche (2 Mädchen und 8 Jungen) mit einem altersbasierten Body-Mass Index (BMI) über der 97th Perzentile (Mittelwert ± Standardabweichung: Alter: 14.6 ± 2.8 Jahre, Körperhöhe: 169.3 ± 11.3 cm, Körpermaße: 99.2 ± 21.7 kg; BMI: 34.2 ± 3.9 kg/m²) wurden rekrutiert und mittels 3DGA-Einheiten untersucht. Eine multivariate Analyse nach Sangeux and Polak wurde verwendet um den repräsentativen Versuch (MRT) aus fünf Versuchen zu ermitteln. Eine durchschnittliche Kurve (AVG) wurde für jeden kinematischen Parameter berechnet. Die Ähnlichkeit der Kurven zwischen den repräsentativen und den gemittelten Versuchen wurde mittels eines linear fit Model (LFM) und mittlerer quadratischer Abweichung (RMSD) berechnet. Zusätzlich wurde die Test-Retest Reliabilität mittels dem Standard Error of Measurement (SEM) kalkuliert.

Ergebnisse: Die Ergebnisse des LFM zeigten eine gute Ähnlichkeit der Kurven zwischen AVG und MRT. Für alle drei Ebenen war der Determinationskoeffizient (R^2) im Durchschnitt über 0,9. Den niedrigsten Wert wies der „foot progression angle“ in der Transversalebene auf (R^2=0,82), während andere Werte wie z.B. der Knie- und Hüftwinkel in der Sagittalebene eine sehr gute Modelanpassung aufwiesen (R^2=1,0). Die mittlere Quadratische Abweichung zwischen AVG und MRT für die Frontal-, Sagittal- und Transversalebene war im Durchschnitt 0.58° ± 0.2°, 1.13 ± 0.56° and 1.28 ± 0.4° für die erste Testung 0.65° ± 0.19°, 1.25° ± 0.72° and 1.34° ± 0.66° für den Retest. Der SEM zwischen Test und Retest war unter 5° für alle kinematischen Parameter. Dies deutet auf eine moderate bis gute Reliabilität hin.

Zusammenfassung: Die Ergebnisse der vorliegenden Studie zeigen die Nützlichkeit der Verwendung des MRT in der klinischen Ganganalyse aufgrund der hohen Ähnlichkeit zu AVG. Der Vorteil von MRT besteht vor allem in der Praktikabilität und im Erhalt von relevanten Informationen. Zudem lässt sich eine hohe Test-Retest Reliabilität feststellen.

Table of Content

1 Introduction

Childhood obesity represents today one of the most serious health issues and public challenges that affect children and adolescents [1]. The body mass index (BMI) is a widely used marker [2] to assess the range of obesity in individuals. A person is considered obese if its BMI is ≥ 30 kg/m² [3]. The lifetime direct medical costs of the obese children relative to non-obese children have been estimated to amount about nineteen thousand (19 000$) USA dollars per child [4].

Impaired quality of life and diminished independence over the course of the day is observed in the individuals suffering from obesity [5]. Besides the increased risk of development of metabolic disorders caused by obesity (type 2 diabetes mellitus, dyslipidemia, impaired glucose tolerance, hypertension and metabolic syndrome) [6], the detrimental effects of obesity can be observed on the musculoskeletal system and might negatively affect the gait mechanics of obese children as well [7, 8]. The major risk is considered to be the development of varus/valgus misalignment of the knee joint which further might lead to serious cartilage damage and degenerative joint diseases [9]. Therefore, the analysis of kinematic and kinetic gait parameters is essential in order to acknowledge the gait misbalance in obese children in detail.

Clinical gait analysis serves as a gold standard tool which makes it possible to estimate and classify more objectively and accurately particular gait disorder or kinematic misbalance between limbs or joints. Three-dimensional gait analysis (3DGA) represents a non-invasive analysis of gait kinematics and allows quantitative assessment of human gait. In most studies, the researchers used to collect several strides and trials during clinical gait analysis.

In clinical practice, the decisions are made based on all the strides collected while the researcher might be more interested in the patient's most representative stride. Therefore, to appropriately analyze and represent the most common person's gait pattern, researchers have used different methods such as averaging the stride curves point/time wise or choosing the one stride visually [10]. Relevant information from the gait pattern might be lost if the averaging method is utilized whereas visual inspection of the most representative stride curve represents a time-consuming job [10]. Therefore, a statistical method was proposed recently in

order to quickly find the most representative gait stride pattern based on multivariate data analysis [10]. The novelty of this approach lies in simplicity to detect outliers i.e., curves which might have been considered to deviate from the representative one markedly.

However, regardless of the implemented measurement protocol to acquire one kinematic data set, the reliability of two consecutive measurements is the next problem that plays a crucial role in 3DGA. In the clinical practice, it is necessary to identify whether the observed change is the result of the actual intervention or simple measurement error [11].

The critical problem occurs when it comes to comparison between data sets (inter-session reliability) where potential measurement errors could mask clinical relevant changes. The consistency of collected kinematic gait data represents a crucial issue due to a variety of factors influencing the position of markers. It is assumed that greater amount of subcutaneous fat tissue in overweight and obese subjects can impair correct identification of anatomical landmarks and therefore can lead to inconsistent marker placement.

As regards to the issues mentioned above, several concerns have arisen which will be structured as follows:

1. The main purpose of this master thesis is to examine if it makes a difference in kinematic measures if one uses an average across five trials (AVG) or the most representative trial (MRT) identified by multivariate data analysis [10].
2. Furthermore, the similarity of the entire waveforms between the average across five trials and the representative trial will be examined [12].
3. The reliability of both methods will be assessed during a test-retest study. The primary aim is to assess if the AVG and MRT will affect test-retest reliability.

1.1 Obesity in the childhood and adolescence

The rapid increase of childhood obesity in the last few decades has raised concern among many health practitioners worldwide [13]. It has been estimated that in 2014 worldwide over 41 million children were affected by obesity or overweight [14]. If this trend continues in the future as well, by 2025 there will be 70 million obese and overweight children globally [14]. Without a proper approach to the overweight and obesity problem, overweight infants and young children will likely maintain higher BMI during childhood, adolescence, and adulthood. Overweight represent the proportion of the children with a BMI greater than one standard deviation (SD) and obesity greater than 2 SDs, taken from World Health Organisation (WHO) growth standard median (Figure 1 and 2) [14]. Furthermore, in Austria, 20% of the girls and 25% of the boys between 7 and 14 years of age are considered overweight, whereas 6 and 9 % are considered obese, respectively [15].

Etiology of obesity in children and adolescents

The most common cause of obesity is an imbalance between energy intake and expenditure [6]. Having poor eating habits combined with low levels of physical activity contributes together to overweight and obesity among children and adolescents [13]. Physical activity regulates energy balance through metabolic and hormonal pathways and has altogether with eating habits the fundamental positive influence on weight control [13]. Other factors which may cause excessive body weight gain include endocrine disorders, genetic syndrome, monogenic obesity, environmental and socio-cultural factors [6].

Consequences of obesity

The consequences of being obese as a child or adolescent are twofold, including both medical and psychosocial co-morbidities [16].

The common medical co-morbidities associated with childhood obesity are listed below [6, 16]:

- Nearly 90% risk of being obese as an adult
- Impaired glucose tolerance and type 2 diabetes mellitus
- Hypertension
- Hyperlipidemia
- Metabolic syndrome

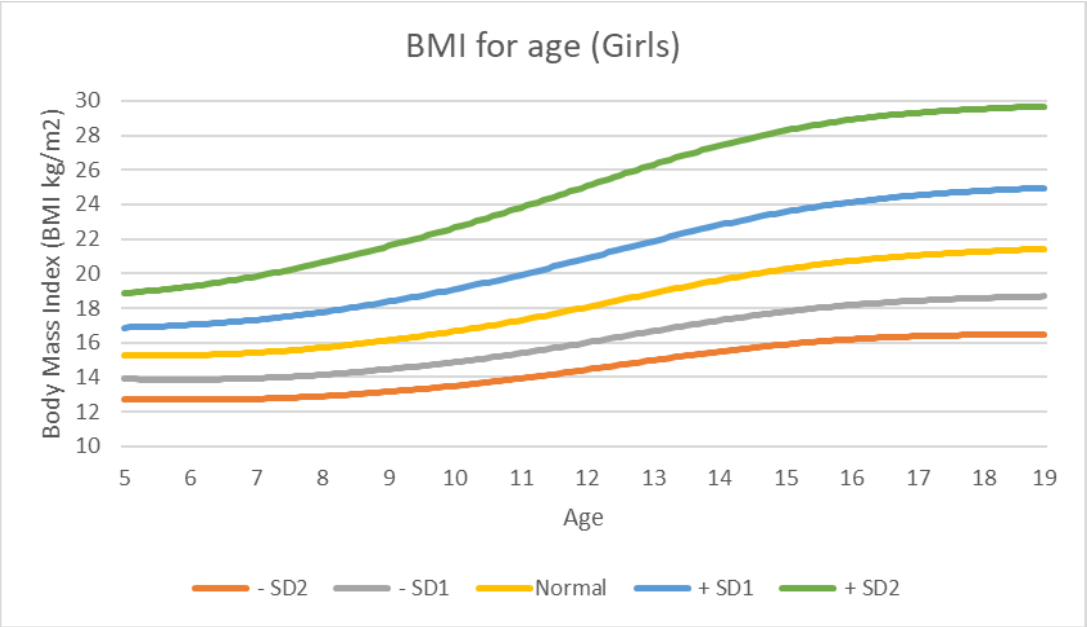

Figure 1 Age-related BMI for the girls. Overweight: >+1SD (equivalent to BMI 25 kg/m^2 at 19 years). Obesity: >+2SD (equivalent to BMI 30 kg/m^2 at 19 years). [17, adapted from WHO].

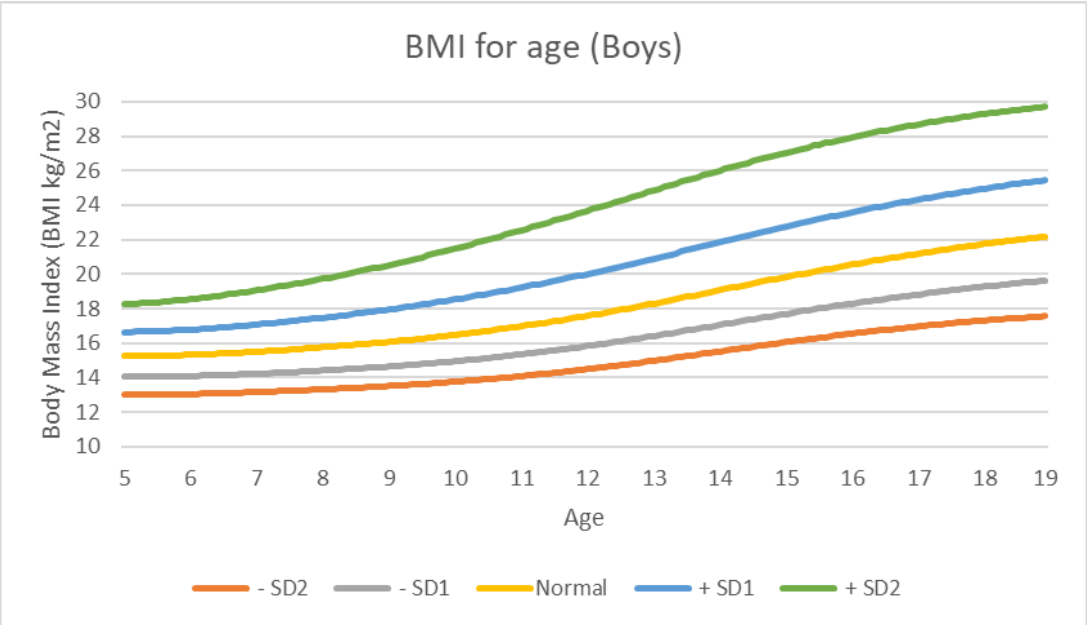

Figure 2 Age-related BMI for the boys. Overweight: >+1SD (equivalent to BMI 25 kg/m^2 at 19 years). Obesity: >+2SD (equivalent to BMI 30 kg/m^2 at 19 years). [17, adapted from WHO].

- Asthma
- Orthopedic disorders:
  Blount disease (Tibia vara)
  Slipped capital femoral epiphysis
- Depression, psychosocial, stigmatization, poor self-esteem

There is strong evidence that the obese children and adolescents suffer from serious health conditions in childhood and have a higher risk to become obese in adulthood, with resulting increased risk of diseases listed above [13]. The continuous increase of childhood obesity worldwide will have dramatic implications in terms of health and economic burden [13].

Treatment

Two main action areas must be recognized and undertaken to get along with increased incidence of childhood obesity: promotion of healthy food intake or eating habits and promotion of physical activity [14]. Caution must be taken with excessive calorie restriction. The goal should be enough energy to keep normal body growth and by avoiding an abnormal increase in body weight in the same time [6]. On the other hand, the physical activity must be a long-term goal, i.e. constituent part of the life style. Similar to adults, children should be physically active most days of the week [18]. They should incorporate at least 60 minutes of moderate to high-intensity physical activity on a daily basis [6, 18].

1.2 Childhood obesity, gait characteristics and musculoskeletal problems

The gait pattern differs between obese and non-obese children and adolescents [19, 20]. The obese children generally tend to walk slower [19] using lower cadence, longer stance period and greater stride width than normal body weight peers [20]. Obese children demonstrate difficulties when attempting to move at speeds other than their preferred walking speed [20]. The kinematic data from several studies have shown that obese children have general tendency to spend a shorter time in single leg support and a tendency toward dual stance [8]. People are walking less as they used to. The children and adolescents aged between 10 and 19 have walked 25% less in 2000's than their peers did in 80's [21, Fig. 3]. Teenage obesity may develop for the reason that every next generation is less physically active than the previous one [21]. If we take all the circumstances of modern society today, we can assume that this trend will continue.

There is no doubt that the walking as a form of physical activity is very important for obese individuals in order to have an optimal weight management program. It is recommended that girls and boys should accumulate between 12 000 and 15 000 steps/day to keep the optimal body weight [22]. Therefore, because of a considerable number of steps per day which transport the ground reaction forces orthogonal throughout the lower body kinetic chain, the correct walking technique and optimal lower body joint alignment during walking are of great importance for obese young individuals [22, 23].

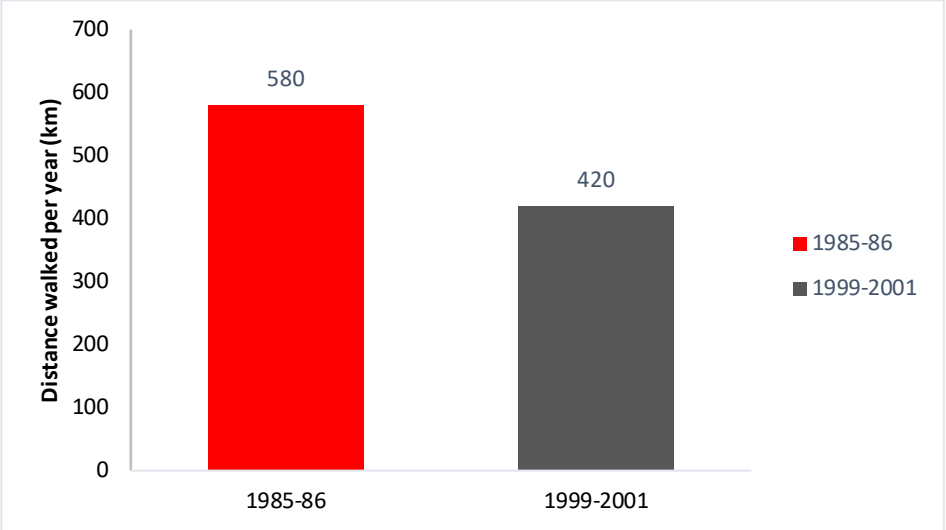

Figure 3 Average distance in kilometer walked per year for children and adolescents between 10-19 of age (data from the National Travel Survey, Department for Transport, UK). Adapted from Kirtley [21, pp. 7].

There is evidence that the overweight children are walking with a significantly lower peak knee flexion angle during early stance [24]. Despite the satisfactory gait adaptation in sagittal plane which allows the overweight children to have similar knee extensor load likewise non-overweight peers, the frontal plane alternation which may lead to increased medial compartment joint loads are present and hard to compensate [24]. Because of such gait kinematic, the occurrence of varus angular deformities of the knee joint and medial compartment osteoarthritis are supported with evidence [7, 8, 24].

The frequent orthopedic problem associated with obese children and adolescents includes Blount disease, slipped capital femoral epiphysis, musculoskeletal pain and osteoarthritis [8, 25].

Blount disease is characterized by inward turning of the lower leg, also known as tibia vara. It has been assumed that the abnormal pressure of excess body weight of obese children causes the trauma at the level of growth plates [25].

Slipped capital femoral epiphysis is manifested as hip pain (or even knee pain) and influences the range of motion (ROM) of the movement [25].

There is increased incidence of musculoskeletal pain associated with obesity in obese children when they were compared with their nonobese counterparts [8]. This can be explained by biomechanical changes in obese children which are the result of a compensatory mechanism in order to support excess body weight [8]. As a result of such mechanical compensations, the postural misalignments may result.

Moreover, it has been demonstrated that obese children have twice the risk of developing obesity when they become adults [26]. Long term persistence of obesity may result in increased likelihood of development of osteoarthritis in the knee joint but not in the hip joint [8]. The study of Harms et al. [27] has demonstrated a significant positive correlation between high BMI and total joint replacements in young adults [8].

2 Clinical gait analysis (CGA)

The objective of gait analysis is to capture and track the kinetic and kinematic data of different body segments [28]. There is evidence supporting the clinical utility of clinical gait analysis [29]. Moreover, the rehabilitation and functional outcome are superior if they were based on gait analysis recommendations [29]. Clinical gait analysis typically includes the following analysis: video analysis, general gait parameters analysis, kinematic analysis, kinetic analysis and surface electromyography (sEMG) [28]. Kinematic parameters are obtained with retro-reflective markers and special cameras connected to corresponding software capable of tracking, processing and storing the motion of the attached markers [28]. The primary kinetic parameter obtained from CGA is the ground reaction force (GRF). The joint torques and powers can then be derived from the kinematic and kinetic parameters [28]. All these parameters enable an easier diagnostic of gait pathology and deviations from optimal gait mechanics. Furthermore, by virtue of CGA the decision-making skills of a treatment provider may be improved [30]. For example, analyzing the kinematic and kinetic data on patient's pre- and post-intervention may provide feedback to clinicians/physicians about the outcome of intervention so they can learn from their mistakes [30-33].

In order to choose an appropriate intervention (or non-intervention) modality, the CGA has to be performed. The reasons to perform CGA is based on the following [34]:

- Differentiate between different disease categories. It is essential that the diagnostic system allows distinction between normal and pathological gait patterns and between different pathological gait characteristics. It is necessary to possess an accurate measurement system and adequate knowledge about characteristics of normal and pathological gait.
- Assessment of the severity, extent or nature of a pathological gait. The measurement system has to be capable of distinguishing clinically important differences between the patients with the same pathological gait characteristics.
- The monitoring of the patient's progress (with or without intervention). The adequate accuracy of clinical gait analysis is necessary to distinguish whether a patient's condition is stagnating, improving or becoming worse.
- Predicting the results of intervention (non-intervention).

2.1 Three-dimensional gait analysis (3DGA)

The 3DGA consists of several cameras which are simultaneously capturing and tracking the images. The captured images are used to reconstruct the trajectory of the point of interests (joints, body segments, etc.) [21]. Compared to 2D gait analysis, any object/point can be tracked with 3DGA as long as it can be seen by at least two cameras [21]. Each camera calibrates the distance and position of a captured marker (volume calibration) separately and makes a relationship to other cameras [21]. The accuracy of 3DGA is typically around ± 0.1% meaning that in the typical gait analysis laboratory (5 meters long) it amounts ± 5 mm [21, 35, 36]. The accuracy is in the most cases limited to the motion artifact of skin-mounted markers [37, 38].

The repeatability represents one of the main problems within 3DGA. There are two major sources of an error occurring during 3DGA. One is related to model calibration and another is related to soft tissue artifacts [34]. Model calibration includes placing a marker on correct anatomical landmarks and the corresponding location of joint centers with respect to the marker location [34]. Lack of appropriate knowledge and sufficient experience regarding anatomical landmarks represent a major contribution to measurement variability due to failure to place a marker accurately by repeating measurements [34]. The experience and training in marker placement are just one side of the coin, the other side represents the lack of appropriate guidelines which are considering placing of markers with regards to certain conditions of patients e.g., excessive subcutaneous tissue in patients which makes palpation of landmarks very difficult [34]. The soft tissue movement represents the second source of measurement error. The different types of soft tissues like skin and muscles possess a certain degree of freedom in relation to bones. The movement of the skin is apparent during walking resulting in the movement of placed markers. This marker displacement provoked by soft tissue characteristics can be processed with optimized algorithms and band pass filters [34]. In the end, the accuracy of a 3DGA critically depends on the experience of the researcher to correctly place the markers [21]. Therefore, the repeatability of 3DGA measurement regarding marker placements will be one of the topics of this master thesis.

3 Quantitative data analysis in 3DGA

Every kinematic variable possesses temporal (time) and spatial (space) characteristics. Walking is a very complex activity i.e., each step has unique temporal and spatial characteristics which are different from previous one. In 3GDA, the fluctuation in the value of a kinematic (e.g., joint angle), kinetic (e.g., ground reaction force), spatio-temporal (e.g., stride interval) or electromyographic measurement is common [39]. Therefore, the need to standardize the gait analysis method in order to make a data comparable between different subjects, studies and research methods were required. One of the most used and simplest methods to make the data comparable is normalization of data sets. Due to the fact that the temporal characteristics of gait vary greatly (intra-subject and inter-subject variability) the time is often normalized to the percentage of the gait cycle [39]. The initial contact to initial contact of the same leg was used as starting and final point of the gait cycle.

The kinematic or kinetic variables in 3DGA can be represented as discrete and continuous variables. A continuous variable can take an infinite number of uncountable values. An example of a continuous variable in 3GDA is a whole-curve analysis of one gait cycle. The advantage of whole-curve analysis is that it does not ignore relevant data of the gait cycle compared to discrete point analysis. For instance, a researcher might be interested in dynamic or development of the entire gait curve and therefore collect important insights about individual gait picture. Furthermore, during the analysis of the entire curve of gait, one could consider the relationship between different stance phases during walking. Nevertheless, the discrete variable is countable and assumes a certain value from the finite data set. In 3DGA discrete variables can represent the initial or terminal point of the gait cycle, ROM, minimal or peak value of the data set, etc. There are some advantages of discrete point analysis over whole-curve analysis allowing the researcher to understand the specific portion of the gait curve. For example, if one is considering the knee injury prevention, it may be useful to understand to which extent the peak landing forces may be harmful to the kneecap of the femur cartilage [40].

During the 3GDA the participants never perform solely one step or one stride but more of them because the researchers are interested in participant's gait pattern which cannot be defined solely in one step or stride. After obtaining data from several steps during 3GDA, the question arises about which step represents best

the subject's gait pattern. In the most studies, this problem was resolved by getting the average data set from the performed gait trials. The next problem that emerged concerning averaging of gait trials is that it can remove the important features from gait data [10, 39]. Therefore, alternative approaches were recommended such as algorithmic calculation of most representative trial, which reflects best the participant's gait pattern [10].

3.1.1 Multivariate data analysis (Sangeux's and Polak's model)

The most representative trial can be calculated for a single kinematic variable or multiple kinematic variables [10]. The MRT should represent ideally a subject's gait pattern across all the variables that form the gait profile [10]. Sangaux and Polak (2015) have proposed the method which can identify the most representative trial (MRT) and outlier across multiple data sets for a single kinematic variable or for several kinematic variables of the gait profile [10, Figure 4]. More precisely the algorithm takes into account several kinematic variables (e.g., pelvic tilt, hip flexion, knee flexion, ankle dorsiflexion and foot progression angle) and identifies the most representative trial which is the same for all kinematic variables used in the calculation. The calculation is based on the concept of centrality with regard to multivariate data which may be further expressed as depth [10]. The deepest value obtained from multivariate data sets represents the most representative value. The at least deepest value corresponds to an outlier.

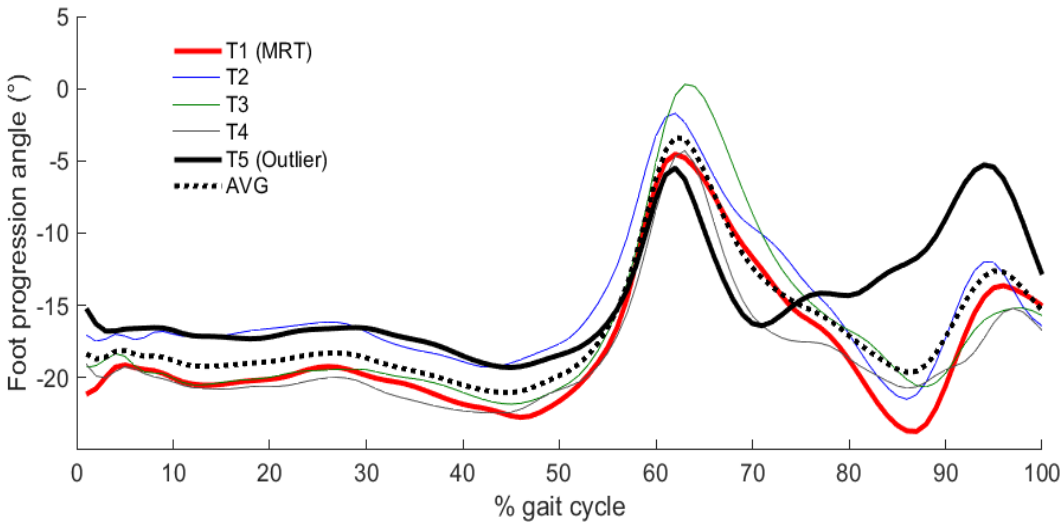

Figure 4 Foot progression angle in the transversal plane of one subject. Data are time normalized to 100% gait cycle. T1-T5 represent five trials. – Most representative trial (MRT) according to Sangaux model, – Outlier, ··· Average (AVG) of five trials.

3.1.2 Waveform Similarity: Linear Fit Method (LFM)

The critical issue in the 3D gait analysis refers to comparison between the data sets i.e., evaluation how much the gait pattern under analysis deviates from reference data [12].

The LFM method represents the linear agreement between two data sets plotted one against the other. More precisely, it estimates how large is the similarity between two plotted curves (Figure 5). The formulas below serve to get three LFM parameters R^2, a_0 and a_1 which are assessing the similarity between the curve under the analysis (*Pa*) and reference curve (*Pref*):

$$a_1 = \frac{\sum_{i=1}^{N}\left(P_{ref}(i)-\overline{P_{ref}}\right) * \left(P_a(i)-\overline{P_a}\right)}{\sum_{i=1}^{N}\left(P_{ref}(i)-\overline{P_{ref}}\right)^2}$$

$$a_0 = \overline{P_a} - a_1 * \overline{P_{ref}}$$

$$R^2 = \frac{\sum_{i=1}^{N}\left(a_0 + a_1 * P_{ref}(i) - \overline{P_a}\right)^2}{\sum_{i=1}^{N}\left(P_a(i)-\overline{P_a}\right)^2}$$

Iosa et al. (2014, p. 2) have explained three LFM parameters as followed [12]:

- *a_1 estimates the mean variation of Pa for every one-unit change in Pref. It hence represents the amplitude scaling factor, that is, the factor for which Pref should be multiplied to match Ya except for a scalar addition.*
- *a_0 predicts this scalar addition (shift), that is, the value of Pa when Pref is equal to 0.*
- *R^2 measures the strength of the linear relationship between Pa and Pref, that is, the percentage of variance in Pa that can be matched by the variance in Pref [12][1].*

[1] In the original article of Iosa et al. (2014) the authors forget to square the numerator for R^2 calculation.

When the perfect similarity between *Pa* and *Pref* exist then the values of LFM parameters are $a_1 = 1$, $a_0 = 0$, $R^2 = 1$.

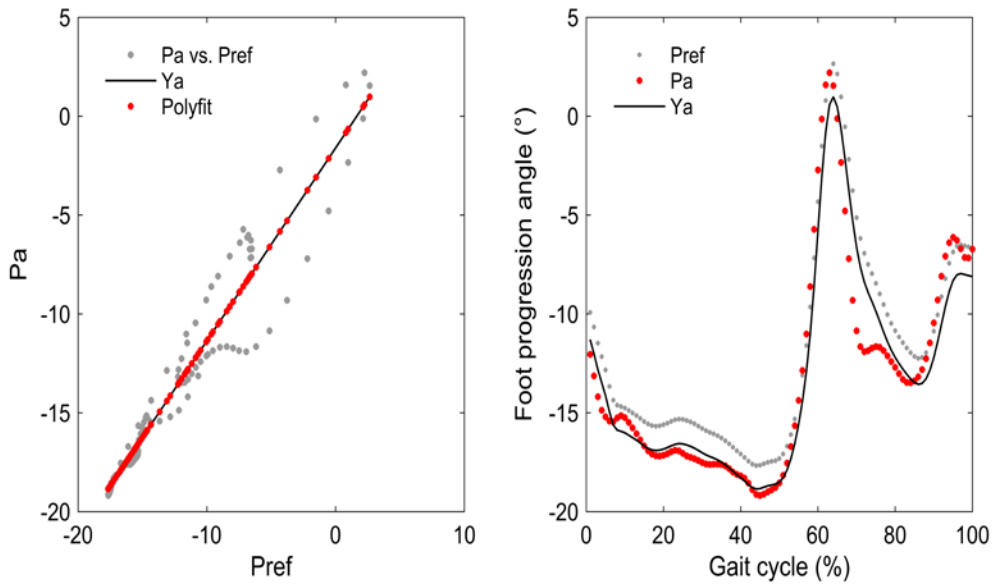

Figure 5 Graphical illustration of LFM for a foot progression angle. Pa – points for the investigated data set Pa and for the reference data set Pref ···. Ya – the linear function which approximates Pa values by means of a linear transformation of values of Pref.

4 Reliability

While the reliability in the literature is considered to represent the consistency of measurements in general, or of an individual's test performance, it can be defined as the absence of measurement error as well [41]. Caution is advised when it comes to interpretation of the reliability because there is no unique reliability assessment tool which might be sufficient to get the full picture about the reliability of data [42]. It is advisable to combine different reliability estimates together [42].

Random and systematic error

The error in statistics generally means the deviation of an observed value from the true value. There are two kinds of these uncertainties or errors: random and systematic [21].

Random error is considered as a "normal" difference arising due to inherent biological or mechanical variation, or inconsistencies in the measurement protocol [41], e.g., the stopwatch may be pressed slightly too early or too late when a runner passed a finishing line [21]. One positive characteristic of random error is that, although it demonstrates the variability of the data it does not affect the mean. This is because each next measurement is just as likely to be above as to be below the mean [21].

In contrast, the systematic error or bias demonstrates a general trend for measurements causes the mean to deviate either positive or negative from its true value [21, 41]. For example, the learning effect might influence that the jump height values in retest are higher than a prior test or that the plastic caliper shows higher pressure (and therefore lower values than actual) at the skinfolds bigger than 40 mm due to greater spring deformation. The negative characteristic of systematic error is the inability to remove it by averaging [21].

Absolute reliability represents individual variability to repeated measurements and it is expressed in actual units of measurement [41, 42].

Relative reliability is the degree to which the measurement keeps its position in the sample over repeated measurements [41, 42].

4.1 Reliability methods

4.1.1 Standard error of measurement (SEM)

The standard error of measurement (SEM) represents the indicator of absolute reliability and it is expressed in the actual units of measurements [41, 42]. Therefore the interpretation of the SEM results is easy where the smaller the SEM value is the greater the reliability [42]. The advantage of SEM is that it is unaffected by the range of measurement [41]. It is expected that if a subject undergoes the same test infinite number of times, it will generate different results from trial to trial due to measurement error. Those measurement errors when plotted would take places above and below the mean. The more reliable the measurement is, the less error dispersion would be around the mean [42].

There are two common ways of calculating the SEM [41]:

- $SEM = SD\sqrt{1 - ICC}$,

  where SD represents the standard deviation of the sample and ICC is the calculated interclass correlation coefficient. The SD in the equation partially disregards the interindividual variation that was used in the calculation of the ICC [41]
- $SEM = \sqrt{\sigma_e^2}$,

  where σ_e^2 is the error variance and equals the mean square error term from an ANOVA [43].

4.1.2 The root-mean-square deviation (RMSD)

The RMSD represents the estimate of total waveform variability between two data sets:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{m}(x_{2.1} - x_{1.i})^2}{m}}$$

For this purpose, the average and the most representative curves for test and retest session were used to compare waveform variability between test and retest for average and the most representative curves as well as between average and the most representative curves for test and retest. Here x_1 and x_2 represent a respected waveform and i=1,….,m is the time index of each time-normalized waveform (0-100% of gait cycle).

5 Test-Retest reliability of kinematic measurements in gait analysis of obese adolescents

5.1 Introduction

Impaired quality of life and diminished independence over the course of the day is observed in patients suffering from gait disorders [8, 44, 45]. Obesity is one of the factors which has detrimental effects on the musculoskeletal system and might negatively affect the gait biomechanics of obese population as well [7, 8, 23]. The major risk is considered to be the development of varus/valgus misalignments of the knee joint [9]. Gait analysis serves as a gold standard tool which allows us to estimate and classify more objectively and accurately particular gait disorder or kinematic misbalance between the left and right lower limb. Three-dimensional gait analysis (3DGA) represents a non-invasive analysis of gait kinematics. However, regardless of the implemented measurement protocol, the reliability plays very important role, as it is necessary to identify whether the observed change is the result of the true intervention or pure measurement error. The size of measurement error (reliability) of 3DGA represents critical issue when it comes to interpretation of gait data [46]. It is assumed that greater amount of subcutaneous fat tissue in overweight and obese subjects can impair correct identification of anatomical landmarks and therefore lead to inconsistent marker placement. However, it is difficult to place the reflective markers in the exact same position on two consecutive testing occasions, especially when it comes to measuring the subjects with excessive subcutaneous fat tissue [47].

In most studies, the researchers used to collect several strides for each individual during clinical gait analysis. In clinical practice, the decisions are made based on all the strides collected while the researcher might be more interested in the patient's most representative stride. In order to find the most representative stride pattern researchers have used different methods such as averaging the gait curves point/time wise or choosing one stride visually or automatically. Important information from the shape of the curve might be lost if the averaging method is utilized whereas visual inspection of the most representative gait curve is a time-consuming job. Therefore, a statistical method was proposed recently in order to easily find the most representative gait stride pattern based on multivariate data analysis [10].

The novelty of this approach lies in its simplicity to detect outliers i.e., curves which might have been considered to deviate from a representative one markedly.

The most prominent issue that will be discussed in this thesis is how to overcome the problems of data consistency in the research environment. Therefore, the two main research questions have been addressed:

- Does it make a difference in kinematic variables if one uses the approach to identify a most representative trial proposed by “Sangeux and Polak” or by using an average over five trials curve?
- Does it make a difference in terms of reliability if we use one or the other approach?

5.2 Methodology

The data used for this study were recorded already elsewhere [48]. However, the study procedure will be described briefly in the next sections.

Participants

A convenience sample of ten participants, two females and eight males with an age-based body mass index (BMI) above 97th percentile [49] were recruited by an outpatient clinic for obese children and adolescents (mean ± SD: age: 14.6 ± 2.8 years, height: 169.3 ± 11.3 cm, body weight: 99.2 ± 21.7 kg; BMI: 34.2 ± 3.9 kg/m2). The participants were eligible for the study if the following inclusion criteria were met:

- Male or female
- Age between 10 and 18
- BMI greater than 97th percentile [50]

The participants were excluded if the following criteria were met:

- Present syndromes associated with obesity (e.g., Prader-Willi syndrome and similar disorders)
- Chronic joint diseases, osteoarthritis surgery
- Neuro-motor diseases

The study received ethical approval from the ethic committee of the Medical University of Vienna (Ethics number: 1445/2013). All subjects were informed about study protocol and before they engaged in the study, they or their legal representatives have signed a written informed consent.

Study design

The subjects visited the biomechanical gait analysis lab on two occasions separated by a minimum of one day between the test and retest sessions (on average, mean ± SD; 3.4 ± 2,0 days). All participants were tested on both test sessions by the same assessor who has one year of experience in 3DGA. During the first test session, the participants walked barefoot on level ground at self-preferred walking speed on a 12-m walkway. The subjects were walking along a 12-m walkway until a minimum five valid trials for both lower body extremities were recorded. The mean walking speed was determined during the first test session using photoelectric sensors. For the retest, the subjects' walking speed was allowed to fluctuate ± 5% from the mean walking speed obtained during the first test session. The kinematic data were obtained for five joints (Figure 6) in all three planes.

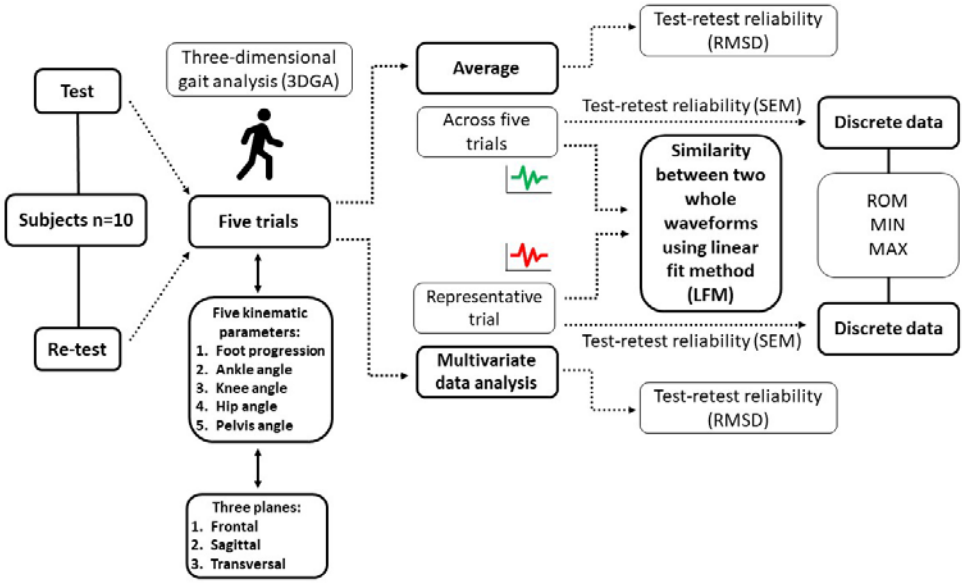

Figure 6 Study design: each participant visited the laboratory for two sessions: test and retest. Five trials were recorded and five kinematic parameters for each plane were exported and time-normalized to 100% gait cycle for further analysis. The data have been averaged across five trials (AVG) and the most representative trial (MRT) has been obtained from multivariate data analysis. Furthermore, the SEM (discrete data) and RMSD (total waveform) analysis were performed between test and retest for AVG and MRT. Additionally, the RMSD and LFM (total waveform) analysis were performed between AVG and MRT for test and retest separately.

Data collection

The Cleveland clinic marker set protocol was used to place the twenty-seven retro reflective spherical markers on anatomical landmarks (Figure 7). Subject's specific

anthropometric measures like leg lengths, pelvis size, knee and ankle widths and joint centers and axes were estimated using anthropometric measuring devices and recorded during static trials. Kinematic data were collected using a motion capture system (Vicon, Oxford, UK) that consisted of two 4.0 and six 1.3 megapixels infrared-cameras recording at a sampling rate of 150 Hz. Zero-lag 4th order low-pass Butterworth filter with a cut-off frequency of 12 Hz was used to filter the raw kinematic data. Data were time-normalized to 100% gait cycle (% gait cycle).

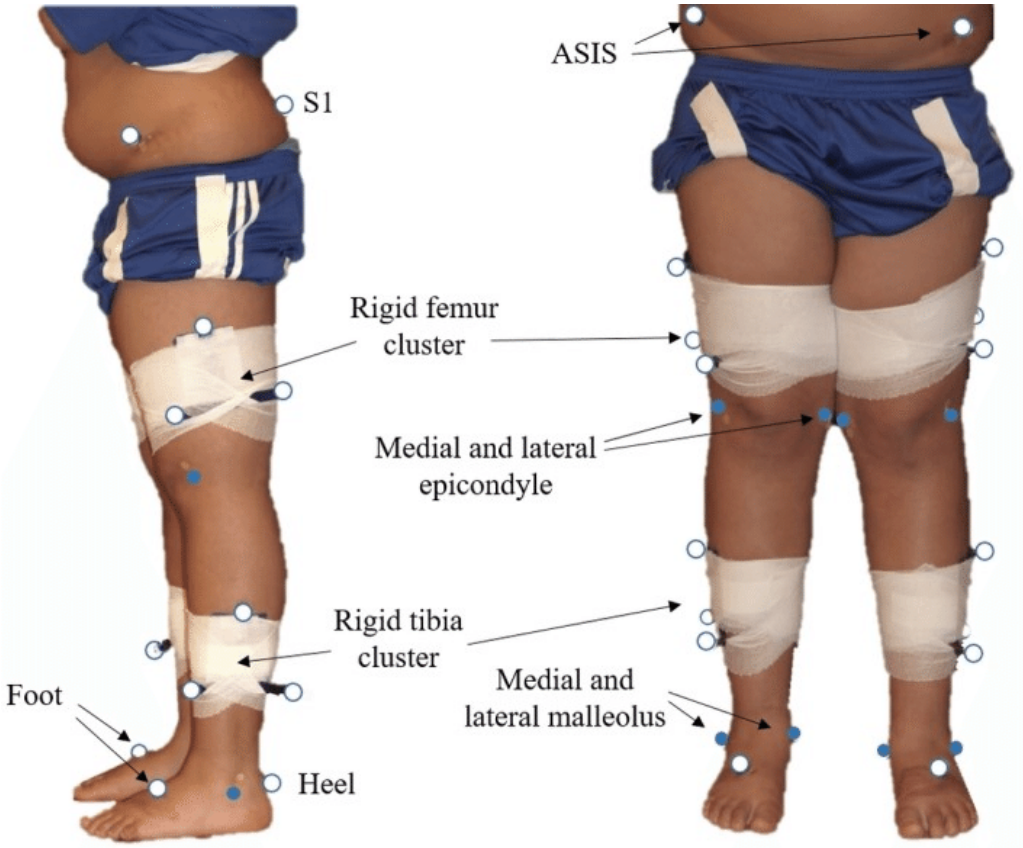

Figure 7 Locations of the retro-reflective markers for the Cleveland Clinic Marker set. Markers attached to the skin as clusters are purely for tracking purpose (Ø 16mm), shaded markers at the knee and ankle are only for anatomical calibration (Ø 9mm). The white markers (Ø 16mm) attached to the foot, heel, the anterior superior iliac spine (ASIS) and the first sacral vertebrae (S1) are for both, anatomical calibration and tracking. Reprinted with permission from Horsak et al. [48].

Data analysis

All kinematic data were analyzed for the left leg arbitrary. Hip, pelvis and knee kinematics were analyzed in all three planes whereas the foot progression angle and ankle angle in transversal and sagittal planes, respectively, resulting in a total

of 11 data sets for every subject per test session. The entire waveforms from each kinematic parameter and for each of the subject's five trials, during the test and retest sessions, were averaged per session.

Furthermore, the simple method to detect the most representative stride across multiple data sets for several kinematic variables of the gait profile, proposed by Sangaux and Polak (2015), was used to identify the most representative trial (MRT) from five gait trials [10].

The entire normalized gait waveforms from all five kinematic parameters were used for every subject and corresponding plane separately in order to identify the most representative trial which is the same for all kinematic variables used in the calculation. For instance, the kinematic data for subject Nr. 1 in sagittal plane included all five joint and multivariate analysis has identified the second trial as the most representative one for all joints in the sagittal plane. This method is sensitive to both shape and position of the curves [10].

The linear fit method (LFM) according to Iosa et al. (2014) was used to estimate strength of linear agreement and shape similarity (R^2) between average across five trials (AVG) and most representative trial (MRT) waveforms as well as amplitude shift and offset (a_1 and a_0, respectively). LFM calculation was performed for every subject and corresponding plane separately.

To estimate the test-retest reliability, the root mean square deviation (RMSD) and standard error of measurement (SEM) were used.

RMSD is used to assess the similarity between two waveforms. The RMSD similarity between test and retest for AVG and MRT as well as the RMSD similarity between AVG and MRT for test and retest were estimated. The RMSD was calculated as follows:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{m}(x_{2.1} - x_{1.i})^2}{m}}$$

where x_1 and x_2 represent the waveforms for test and re-test, respectively, and i….m are the index numbers of the data points in the time series.

Discrete local minima, maxima and range of motion (ROM) were identified and exported for further analysis of the SEM. The SEM for each discrete parameter between the test and retest for AVG and MRT was analyzed. The mean square

error term from an ANOVA was used to calculate the SEM. The SEM represents the consistency of scores within an individual.

$$SEM = SD\sqrt{1 - ICC}$$

Statistical analysis

The Kolmogorov-Smirnov test was performed to estimate distribution normality. Furthermore, a two-sample *t*-test was used to determine if AVG and MRT sample means for test and retest are equal or if the inherited difference between aforementioned samples comes from pure systematic error. All reliability and statistical analyses were performed in Matlab (v. R2012a, The Mathworks, Natick, MA). Matlab scripts [51, 52] and Excel spreadsheet [53] were used to calculate the most representative trial according to Sangeux and Polak [10]

5.3 Results

Waveform similarity between AVG and MRT for test and retest

The average LFM values of the kinematic parameter related to the planes are reported in Table 1. The LFM parameter R^2 demonstrated the lowest reliability in transversal planes for test and retest, 0.94 and 0,91 respectively, whereas the sagittal plane for test and retest has shown the highest reliability, 0.96 and 0.97 respectively. The LFM parameter (a_0) averaged across individual planes ranged from 0° to - 0.39°, showing the highest amplitude shift for test and retest values in the transversal plane, - 0.21° and - 0.39° respectively, whereas the remaining planes demonstrated hardly any amplitude shift (0° to – 0.06°).

Table 1 The LFM parameters averaged across related planes (F, S and T standing for the frontal, sagittal and transversal plane, respectively). The mean and standard deviation are presented for both tests.

	Test			**Retest**		
LFM	F	S	T	F	S	T
a1	0,99 (0,07)	1 (0,06)	1,02 (0,11)	0,98 (0,07)	1,02 (0,08)	0,98 (0,18)
a0	0 (0,79)	0 (0,77)	-0,21 (1,51)	-0,02 (0,97)	-0,06 (1,05)	-0,39 (1,65)
R	0,95 (0,08)	0,96 (0,10)	0,94 (0,07)	0,95 (0,07)	0,97 (0,06)	0,91 (0,16)

In the Figure 8 are presented five trials of one subject with AVG and MRT for each joint in the corresponding plane.

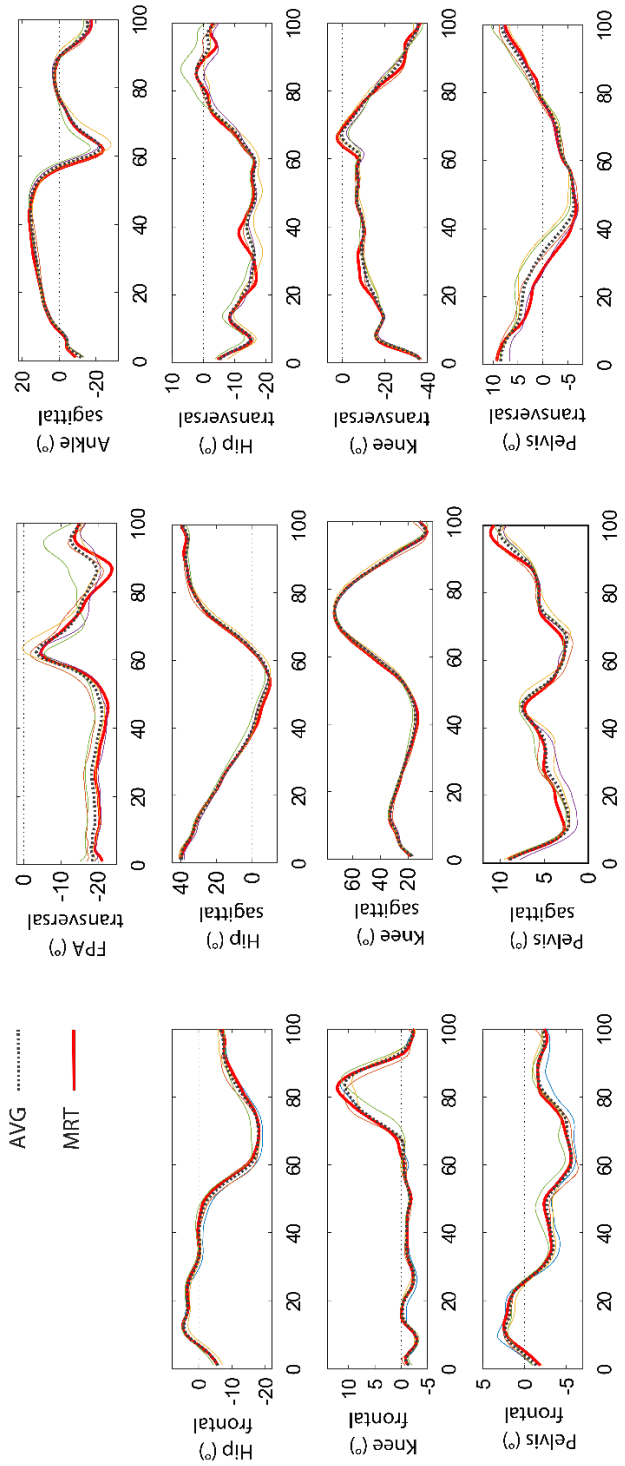

Figure 8 Kinematic waveforms of one subject with their corresponding trials, AVG and MRT. Each graph contains five trials (designated with the different colors), average curve across five trials (··· AVG) and most representative trial (– MRT). Following five joints with their corresponding planes are presented: Ankle angle in sagittal plane, Foot progression angle (FPA) in transversal plane and Knee, Hip and Pelvis angle in all three planes.

Table 2 Mean ± SD of the linear fit method (LFM) between averaged across five trials (AVG) and most representative trial (MRT) for the test and retest, respectively. Data are presented for every joint and corresponding plane.

Kinematic parameter	Planes	LFM	Test	Retest
Foot	Transversal	**a1** (SD)	0,95 (0,14)	0,97 (0,29)
		a0 (SD)	-0,56 (2,58)	-0,98 (3,07)
		R² (SD)	0,92 (0,10)	**0,82 (0,27**)
Ankle	Sagittal	**a1** (SD)	0,98 (0,06)	1,01 (0,04)
		a0 (SD)	0,14 (0,69)	0,42 (1,18)
		R² (SD)	0,98 (0,01)	0,97 (0,02)
Knee	Frontal	**a1** (SD)	1,02 (0,06)	1,01 (0,05)
		a0 (SD)	0 (0,31)	0,1 (0,29)
		R² (SD)	0,99 (0,01)	0,99 (0,01)
	Sagittal	**a1** (SD)	1,01 (0,03)	1 (0,04)
		a0 (SD)	-0,08 (0,54)	-0,06 (0,69)
		R² (SD)	**1 (0)**	**1 (0)**
	Transversal	**a1** (SD)	1 (0,08)	0,97 (0,06)
		a0 (SD)	0,11 (0,90)	-0,28 (0,66)
		R² (SD)	0,92 (0,07)	0,93 (0,03)
Hip	Frontal	**a1** (SD)	1,02 (0,05)	0,96 (0,06)
		a0 (SD)	0,01 (0,39)	-0,14 (0,35)
		R² (SD)	0,96 (0,07)	0,96 (0,03)
	Sagittal	**a1** (SD)	1,01 (0,03)	1 (0,02)
		a0 (SD)	-0,07 (0,90)	-0,04 (0,90)
		R² (SD)	0,99 (0)	**1 (0)**
	Transversal	**a1** (SD)	1,02 (0,06)	1 (0,08)
		a0 (SD)	0,37 (1,06)	-0,41 (0,83)
		R² (SD)	0,97 (0,02)	0,96 (0,02)
Pelvis	Frontal	**a1** (SD)	0,94 (0,06)	0,97 (0,09)
		a0 (SD)	-0,02 (0,31)	-0,02 (0,37)
		R² (SD)	0,92 (0,12)	0,9 (0,1)
	Sagittal	**a1** (SD)	0,99 (0,11)	1,06 (0,14)
		a0 (SD)	0,03 (0,96)	-0,56 (1,25)
		R² (SD)	0,86 (0,16)	0,9 (0,1)
	Transversal	**a1** (SD)	1,1 (0,09)	0,95 (0,21)
		a0 (SD)	-0,77 (0,62)	0,11 (0,74)
		R² (SD)	0,94 (0,07)	0,91 (0,15)

The LFM parameters averaged across all subjects for test and retest are shown in Table 2. The LFM parameter R^2 ranged from 0.82 to 1, demonstrating the lowest reliability for the foot progression angle during retest (0.82, see Table 2). The highest reliability from LFM values was observed in sagittal plane from the knee and hip joints (1) for test and retest, respectively, (Table 2). The LFM parameter *a1* was above 0.92 for all joints, planes and both tests, whereas *a0* ranged between – 0.98° to 0.42°. The largest amplitude shift was observed during retest session in a transversal plane from foot progression angle joint (– 0.98°) and sagittal plane from ankle angle (0.42°).

The RMSD values between the test and retest for AVG and MRT as well as between the AVG and MRT for test and retest for all three planes separately are shown in Figure 9 (for detailed data see Appendix). The highest RMSD was observed in the transversal plane (3.86° ± 2.41°, 3.91 ± 2.2°, 1.28° ± 0.4° and 1.34° ± 0.66°) and the lowest in the frontal plane (2.2° ± 1.04°, 2.33° ± 0.97°, 0.58° ± 0.2°, 0.65° ± 0.19°) in all four analysis protocols.

The RMSD values between AVG and MRT for the test (0.58° - 1.28°) and retest (0.65° - 1.34°) have generally lower values compared to RMSD values between test and retest for AVG (2.2° - 3.86°) and MRT (2.33° - 3.91°) indicating better similarity between the average curve (AVG) and most representative curve (MRT) in both tests.

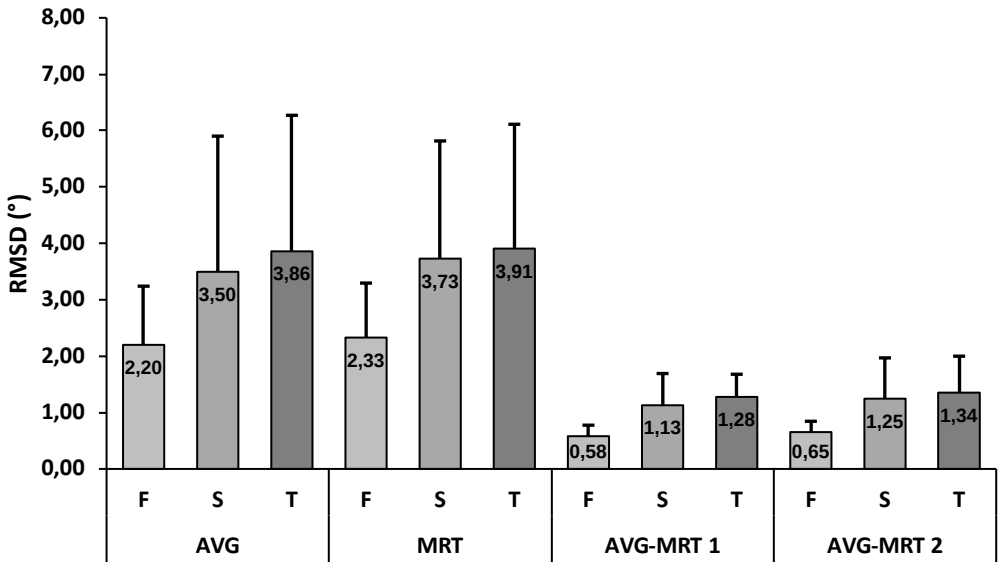

Figure 9 The root mean deviation (RMSD) relative to the kinematic plane. The data represents RMSD between test and retest (left) for average curves (AVG) and most representative trials (MRT) and between AVG and MRT (right) for test and retest, respectively.

Table 3 summarizes the RMSD values relative to kinematic parameter between the test and retest as well as between the AVG and MRT. The lowest values between the tests were found in the pelvis in the frontal plane (1.52° ± 0.6° and 1.76° ± 0.58°, for AVG and MRT respectively) whereas the highest were found in the knee in the transversal plane (5.08° ± 2.53° and 5.36° ± 2.47°, for AVG and MRT respectively). The same trend regarding the lowest RMSD value was observed in the frontal pelvic plane when the comparison was made between AVG and MRT for both tests (0.43° ± 0.12° and 0.56° ± 0.19°, for test and retest, respectively). In contrast, the highest RMSD values were found in knee sagittal plane (1.65° ± 0.65°) when it comes to comparison between AVG and MRT for test and in foot progression angle in a transversal plane for a retest (1.74° ± 1.03°).

Table 3 The root mean deviation (RMSD) for the individual kinematic joints. The data represents RMSD between test and retest (left) for average curves (AVG) and most representative trials (MRT) and between AVG and MRT (right) for test and retest, respectively.

	RMSD (mean ± SD)			
	AVG	**MRT**	**AVG-MRT1**	**AVG-MRT2**
FOOT transversal	3,14 (1,53)	3,36 (1,25)	1,49 (0,51)	1,74 (1,03)
ANKLE sagittal	3,13 (2,48)	2,87 (1,40)	1,25 (0,32)	1,58 (1,07)
HIP frontal	2,88 (1,07)	2,77 (1,07)	0,67 (0,18)	0,65 (0,11)
HIP sagittal	4,69 (3,05)	5,05 (2,58)	1,03 (0,36)	1,03 (0,30)
HIP transversal	4,98 (3,56)	5,37 (3,53)	1,13 (0,21)	1,17 (0,26)
KNEE frontal	2,20 (0,98)	2,46 (0,97)	0,64 (0,19)	0,75 (0,22)
KNEE sagittal	3,09 (1,45)	3,73 (1,53)	1,65 (0,65)	1,61 (0,61)
KNEE transversal	5,08 (2,53)	5,36 (2,47)	1,34 (0,29)	1,55 (0,42)
PELVIS frontal	1,52 (0,60)	1,76 (0,58)	0,43 (0,12)	0,56 (0,19)
PELVIS sagittal	3,07 (2,34)	3,26 (2,22)	0,59 (0,26)	0,76 (0,29)
PELVIS transversal	2,53 (1,65)	2,21 (1,09)	1,14 (0,46)	0,91 (0,29)
Mean(SD)	3,3 (1,15)	3,47 (1,27)	1,03 (0,4)	1,12 (0,43)

The standard error of measurement (SEM) between the test and retest for AVG and MRT, similar as RMSD, demonstrated the same trend concerning measurement error relative to the planes (Figure 10). The discrete kinematic parameters in the transversal plane have demonstrated the greatest SEM regardless of analysis protocol (2.07° ± 0.5° - 3.57° ± 0.92°, see Appendix), whereas the values for the discrete kinematic parameters in the frontal plane have shown the lowest values (1,48° ± 0.56° - 1.95° ± 0.44°, see Appendix).

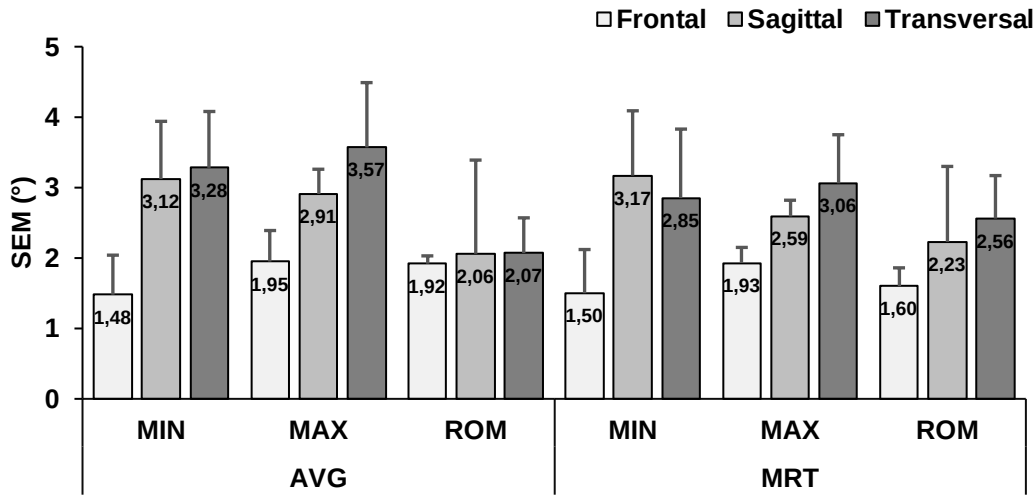

Figure 10 The standard error of measurement (SEM) averaged across the kinematic planes. The data represents the SEM between test and retest for the AVG and the MRT for the discrete kinematic parameter (MIN, MAX and ROM).

The SEM values between test and retest averaged across each discrete kinematic parameter (MIN, MAX and ROM) separately, are shown in Figure 11. The ROM demonstrated the lowest SEM (2.03° ± 0.78 ° - 2.18° ± 0.79°, for AVG and MRT, respectively) while the MIN and MAX were above 2.5° for both curves (AVG and MRT).

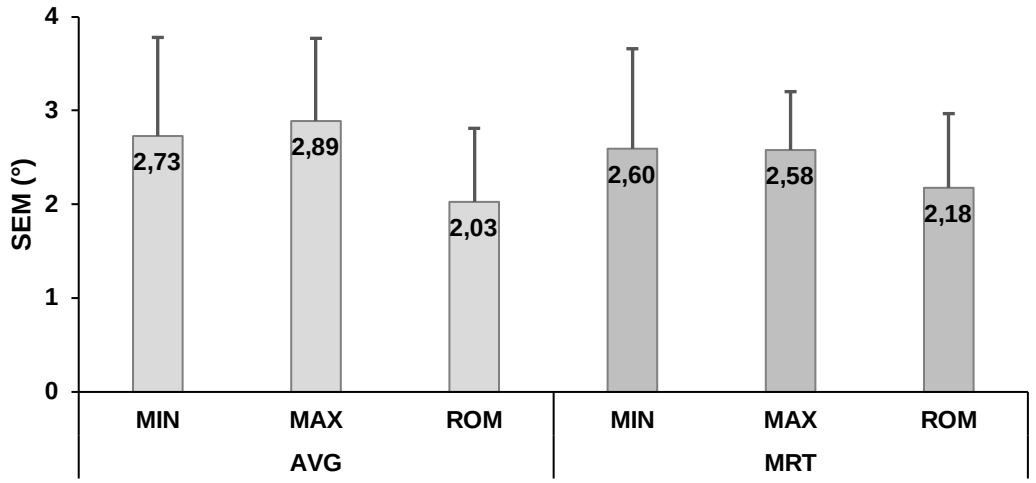

Figure 11 The SEM values of discrete kinematic parameters averaged across all the joints. The figure shows the data of the test compared to the retest for AVG and MRT.

Table 4 The SEM values of discrete kinematic data for each kinematic joint when compared test to retest for AVG and MRT.

	SEM					
	AVG			MRT		
	MIN	MAX	ROM	MIN	MAX	ROM
FOOT transversal	3,6	4,57	1,5	2,35	2,79	3,3
ANKLE sagittal	3,59	3	2,29	3,5	2,41	2,27
HIP frontal	1,94	2,43	1,87	2,07	2,01	1,37
HIP sagittal	4,01	3,33	1,59	4,14	2,77	2,33
HIP transversal	3,65	3,37	2,48	3,45	3,48	2,14
KNEE frontal	1,64	1,85	2,04	1,59	2,1	1,89
KNEE sagittal	2,23	2,49	3,77	1,95	2,37	3,47
KNEE transversal	3,78	3,94	2,51	3,86	3,75	2,8
PELVIS frontal	0,86	1,57	1,85	0,83	1,68	1,53
PELVIS sagittal	2,66	2,84	0,6	3,09	2,81	0,87
PELVIS transversal	2,08	2,41	1,79	1,73	2,23	1,99
Mean (SD)	2,73 (1,05)	2,89 (0,88)	2,03 (0,78)	2,6 (1,07)	2,58 (0,62)	2,18 (0,79)

In Table 4 are outlined the SEM values of discrete kinematic data for each kinematic joint when compared test and retest for AVG and MRT. The SEM values showed less than 5° error of measurements for all joints in discrete kinematic data. The pelvis in the frontal plane shows the lowest SEM values in all discrete parameters ranging from 0.83° to 1.85° for MIN and ROM. The MIN parameter of

the hip joint in sagittal plane demonstrated the highest SEM value 4.01° and 4.14°, for AVG and MRT respectively. The lowest SEM was observed in ROM parameter of the pelvis joint in the sagittal plane, 0.6° and 0.87° for AVG and MRT, respectively.

The two-sample Kolmogorov-Smirnov test revealed the normal distribution of discrete kinematic data when analyzed across all subjects per joint and plane for average gait curves (AVG) and most representative curves (MRT). In addition, the normal distribution was examined with the same test for the SEM values in Table 4 for each discrete data set (MIN, MAX and ROM) and test has shown normal distributed of SEM data sets.

Furthermore, a *t*-test was utilized to quantify if two sample means are equal or inherent systematic error between samples exist. The p-values for each kinematic parameter were outlined in Table 5. There was no statistically significant difference between sample means (p-values ranging from 0.5 to 1.0).

Table 5 Values from the t-test performed between AVG and MRT discrete data sets for test and retest separately. Data indicating that the variance between two samples for each kinematic variable is statistically insignificant.

		Test			**Retest**		
		MIN	***MAX***	***ROM***	***MIN***	***MAX***	***ROM***
Foot	T	1,0	0,81	0,76	0,99	0,65	0,62
Ankle	S	0,93	0,96	0,94	0,94	0,63	0,7
Hip	F	0,97	0,89	0,86	0,98	0,92	0,94
	S	0,91	0,9	0,84	0,88	0,94	0,87
	T	0, 98	0,63	0,39	0,86	0,98	0,78
Knee	F	0, 86	0,89	0,84	0,95	0,72	0,82
	S	0,84	0,87	0,89	0,85	0,98	0,92
	T	0,94	0,8	0,65	0,88	0,85	0,62
Pelvis	F	0,96	0,87	0,9	0,99	0,83	0,91
	S	0,96	0,86	0,63	0,8	0,86	0,44
	T	0,50	0,93	0,5	0,85	0,81	0,76

Furthermore, the SEM sample means between AVG and MRT discrete parameter across all joints were outlined with the *p*-values from the *t*-test in Table 6. The sample means did not differ significantly showing the *p*-values of 0.77, 0.35 and 0.66 for MIN, MAX and ROM, respectively.

Table 6 p-values from t-test performed between AVG's and MRT's SEM

	AVG-MRT		
	MIN	**MAX**	**ROM**
p Value	0,77	0,35	0,66

5.4 Discussion

The aim of this study was to assess and examine the similarity between the AVG and MRT. The MRT is a novel method proposed by Sangeux and Polak (2015) which choose the most representative gait curve among repeated trials [10]. Some crucial decisions could be made in clinical practice due to quantification of gait pattern deviation under analysis from reference data [12, 32]. In order to make gait data comparable between different subjects or between pathological and healthy gait pattern, the averaging method point/time wise was mostly utilized in research practice. Nevertheless, the clinician might be interested in one representative curve which reflects the most relevant characteristics of an individual gait pattern [10, 54]. The averaging across multiple trials might lead to the removal of important features from the shape of gait curves such as peak amplitude or rate of amplitude development [10].

In general, the results indicated a good similarity for entire kinematic waveform between the AVG and MRT curves justifying the use of MRT in clinical practice. In the Figures 12,13 and 14 are presented AVG and MRT among other relevant gait curves for the transversal, frontal and sagittal planes, respectively. The left picture includes, besides five gait trials, the AVG as well, whereas the right picture includes only AVG and MRT gait curves.

Overall the similarity between AVG and MRT curves based on LFM analysis showed the highest linear strength relationship (R^2) for the sagittal (0.95-0.97) and frontal planes (0.95) compared to transversal plane (0.91-0.94) and lower offset (a_0) ranging from 0° to -0.06° and from -0.21° to -0.39° for sagittal/frontal and transversal plane, respectively. This data are in accordance with McGinley et al. [55] and Meldrum et al. [46], who reported the lower reliability and consistency of overall kinematic data in the transversal plane when compared to frontal and sagittal planes. Moreover, besides the lower curve similarity between AVG and MRT for the foot progression angle (R^2 = 0.92/0.82 and a_0 = -0.56°/-0.98° for test and retest, respectively) it has been found the lowest similarity between relevant curves for pelvic tilt (0.86 - 0.94; a_0 = 0.11° - -0.77°) when compared to all other

joints. The lower reliability of the pelvic tilt kinematic has been observed in other studies as well [11]. The overall lower waveform similarity of foot progression angle (FPA) could be partially explained by the great inter-subject variability observed in our study (R^2 was ranging from 0.1 to 0.94 and a_0 ranging from 3.09° to – 7.9°).

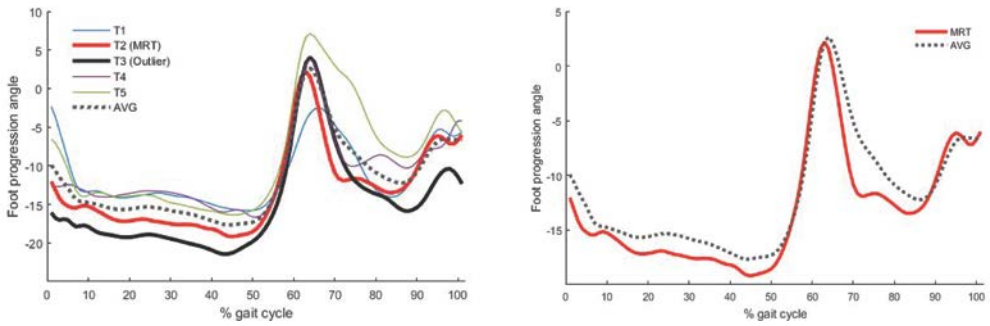

Figure 12 Foot progression angle in the transversal plane of one subject. Left are all five trials including an average of five trials data set, whereas the right is plotted only most representative trial (MRT) and average (AVG) of five trials. Data are time normalized to 100% gait cycle. T1-T5 represent five trials, respectively. – Most representative trial (MRT) according to Sangaux model, – Outlier, ··· Average (AVG) of five trials.

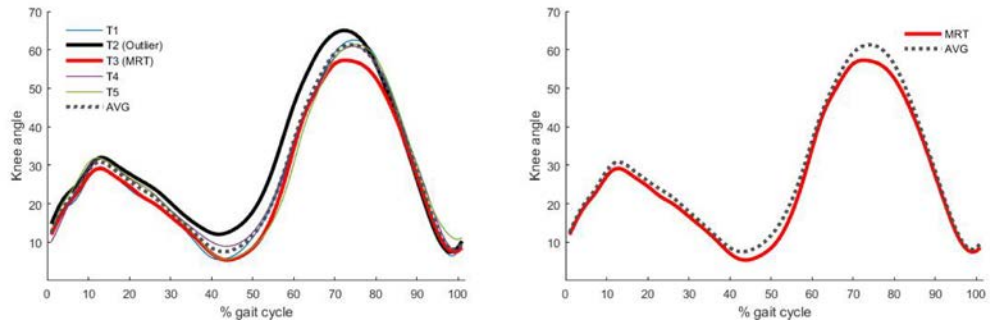

Figure 13 Knee angle in the sagittal plane of one subject. Left are all five trials including an average of five trials data set, whereas the right is plotted only most representative trial (MRT) and average (AVG) of five trials. Data are time normalized to 100% gait cycle. T1-T5 represent five trials, respectively. – Most representative trial (MRT) according to Sangaux model, – Outlier, ··· Average (AVG) of five trials.

It is well documented that the alternation of the FPA during walking modifies the gait pattern in order to reduce knee joint loading, namely the knee adduction moment [56-58]. The lower LFM values in some subjects indicate large inter-trial variability given that average curve can significantly deviate from the most representative curve. The large inter-subject variability in LFM values was

observed in pelvic tilt correspondingly (R^2 was ranging from 0.51 to 0.99 and a_0 ranging from 1.87° to – 3.42°). The wobbling mass of subcutaneous tissue, especially at the pelvis and abdominal region, causes a lot of movement artifacts which are affecting anterior-posterior tilt of the pelvis [48]. Also, a large amount of subcutaneous tissue makes it difficult to identify anatomical landmarks correctly and thus preventing consistency in marker placements consequently provoking uncertainty in measurement [34, 59, 60].

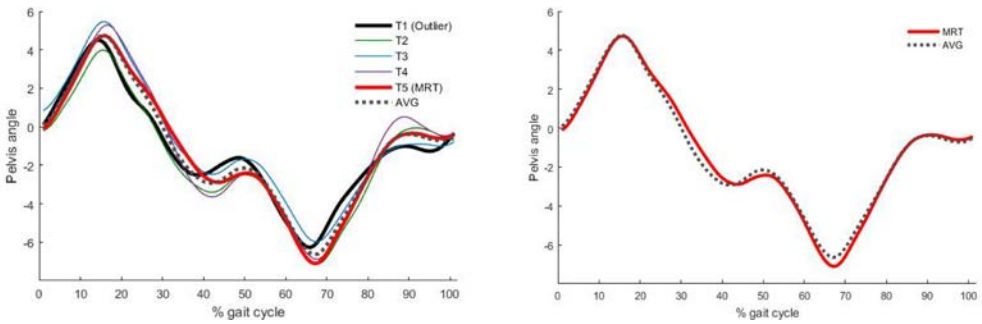

Figure 14 Pelvis angle in the frontal plane of one subject. Left are all five trials including an average of five trials data set, whereas the right is plotted only most representative trial (MRT) and average (AVG) of five trials. Data are time normalized to 100% gait cycle. T1-T5 represent five trials, respectively. – Most representative trial (MRT) according to Sangaux model, – Outlier, ··· Average (AVG) of five trials.

The root mean square deviation (RMSD) showed as well, that kinematic parameters in the transversal plane demonstrated the greater difference compared to a frontal and sagittal plane, regardless of implemented comparisons. We observed uncertainty in the transversal plane when compared the test to retest or average curve to the most representative one. The absolute values of RMSD when compared AVG to MRT for both test were markedly low (<1.74°) indicating that the similarity between AVG and MRT method is acceptable, especially for the frontal and sagittal planes. Average RMSD across all joints between AVG and MRT was on average 1.03° ± 0.4° and 1.12° ± 0.43° for test and retest, respectively, demonstrating the acceptable difference (< 2° on average).

The averaged SEM values across all joints were in acceptable range showing the average error of SEM for all discrete parameters below 2.89° implying good reliability between the test and retest curves. McGinley et al.[55] stated in their systematic review that errors between 2-5° represent a reasonable indicator that the difference between the gait curves is negligible and could be considered as acceptable for further analysis. Hip joint in the sagittal plane, Foot progression

angle (FPA) and knee joint in the sagittal plane showed noticeably higher SEM values for the test (4.01°, 4.57° and 3.77° for MIN, MAX and ROM, respectively). In addition, the hip joint in the sagittal plane, and knee joint in the transversal and sagittal plane showed the highest SEM values for retest (4.14°, 3.75° and 3.47° for MIN, MAX and ROM, respectively) compared to all other joints for test (0.86°-3.78°, 1.57°-3.94° and 0.6°-2.51° for MIN, MAX and ROM, respectively) and retest (0.83°-3.86°, 1.68°-3.48° and 0.87°-3.3° for MIN, MAX and ROM, respectively).

Our results suggest the utility of the method proposed by Sangeux and Polak who defined the most representative trial (MRT) based on the notion of depth, where the deepest curve is the equivalent to the median for univariate data [10]. Moreover, the linear fit method (LFM) showed good linear strength relationship in all planes and joints (R^2>0.82). Iosa et al. (2014) outlined three limitations of the LFM which should be considered when implementing LFM. 1) When $R^2 < 0.5$ the relationship between curves can be only partially described by LFM, 2) difficulty to define which data set is referent one and 3) bias as a result of phase shift (mainly shift along the horizontal gait cycle axis). None of these limitations have been encountered due to lowest estimated R^2 of 0.82, clear differentiation between reference data set (average curve) and data set under investigation (MRT) and normalized data set which revoke phase shift. RMSD and SEM along with the LFM demonstrated equally good similarity and reliability when the average curves were compared to the most representative trial relative to the plane or kinematic joints. The overall RMSD and SEM values were on average around or below 2° suggesting that our results are in accordance with previously published studies [46, 55].

6 Conclusion

This study investigates the usefulness of most representative trial (MRT) in clinical gait analysis and the justification to use a most representative curve in clinical practice due to the retention of relevant features of curve shape. There is a fair indication that the MRT could be used in clinical gait analysis due to good similarity to average curve across multiple trials (AVG) as well as lower RMSD and SEM values between this two curve models. The possible advantage of MRT lies in its simplicity and retention of relevant data/shape information. Furthermore, the considerable lower values of RMSD between the AVG and MRT for the test and retest indicate the slight overall difference between total waveforms. Data suggest that the test-retest reliability for all discrete kinematic parameters was moderate to good.

Literature

[1] K. K. Davison, J. Falbe, E. M. Taveras, S. Gortmaker, M. Kulldorff, M. Perkins, R. E. Blaine, R. L. Franckle, C. Ganter, J. Woo Baidal, J.-A. Kwass, J. Buszkiewicz, L. Smith, and T. Land, "Evaluation Overview for the Massachusetts Childhood Obesity Research Demonstration (MA-CORD) Project," *Childhood Obesity,* vol. 11, pp. 23-36, 2015.

[2] A. C. Buchholz and J. M. Bugaresti, "A review of body mass index and waist circumference as markers of obesity and coronary heart disease risk in persons with chronic spinal cord injury," *Spinal Cord,* vol. 43, pp. 513-518, 2005.

[3] American College of Sports Medicine, *ACSM's Guidelines for Exercise Testing and Prescription*: Wolters Kluwer Health, 2013.

[4] E. A. Finkelstein, W. C. K. Graham, and R. Malhotra, "Lifetime Direct Medical Costs of Childhood Obesity," *Pediatrics,* 2014.

[5] J. B. Schwimmer, T. M. Burwinkle, and J. W. Varni, "Health-related quality of life of severely obese children and adolescents," *JAMA,* vol. 289, pp. 1813-1819, 2003.

[6] A. Seth and R. Sharma, "Childhood Obesity," *The Indian Journal of Pediatrics,* vol. 80, pp. 309-317, 2013.

[7] A. L. De Sá Pinto, P. M. De Barros Holanda, A. S. Radu, S. M. F. Villares, and F. R. Lima, "Musculoskeletal findings in obese children," *Journal of Paediatrics and Child Health,* vol. 42, pp. 341-344, 2006.

[8] G. Chan and C. T. Chen, "Musculoskeletal effects of obesity," *Current Opinion in Pediatrics,* vol. 21, pp. 65-70, 2009.

[9] B. Horsak, D. Artner, A. Baca, B. Pobatschnig, S. Greber-Platzer, S. Nehrer, and B. Wondrasch, "The effects of a strength and neuromuscular exercise programme for the lower extremity on knee load, pain and function in obese children and adolescents: study protocol for a randomised controlled trial," *Trials,* vol. 16, p. 586, 2015.

[10] M. Sangeux and J. Polak, "A simple method to choose the most representative stride and detect outliers," *Gait & Posture,* vol. 41, pp. 726-730, 2015.

[11] M. H. Schwartz, J. P. Trost, and R. A. Wervey, "Measurement and management of errors in quantitative gait data," *Gait & Posture,* vol. 20, pp. 196-203, 2004.

[12] M. Iosa, A. Cereatti, A. Merlo, I. Campanini, S. Paolucci, and A. Cappozzo, "Assessment of Waveform Similarity in Clinical Gait Data: The Linear Fit Method," *BioMed Research International,* vol. 2014, p. 7, 2014.

[13] World Health Organization, "Adolescent obesity and related behaviours: trends and inequalities in the WHO European Region, 2002–2014," 2017.

[14] World Health Organization. (2016). *Commission on Ending Childhood Obesity*.

[15] E. Ibrahim, "Österreichischer Ernährungsbericht 2012," vol. Auflage 1, 2012.

[16] E. R. Pulgarón, "Childhood Obesity: A Review of Increased Risk for Physical and Psychological Co-morbidities," *Clinical therapeutics,* vol. 35, pp. A18-A32, 2013.

[17] World Health Organization. (2017). *Growth reference 5-19 years*

[18] F. D. Stoler, "Childhood Overweight & Obesity," *ACSM Sports Medicine Basics,* 2016.

[19] L. Conti, M. G. Benedetti, S. Sergi, A. Di Gioia, L. Berti, F. Catani, and S. Giannini, "Gait abnormalities in obese people and knee osteoarthritis," *Gait & Posture,* vol. 30, pp. S40-S41, 2009.
[20] A. P. Hills and A. W. Parker, "Gait characteristics of obese children," *Archives of Physical Medicine and Rehabilitation,* vol. 72, pp. 403-407, 1991.
[21] C. Kirtley, *Clinical Gait Analysis: Theory and Practice*: Elsevier, 2006.
[22] S. P. Shultz, R. C. Browning, Y. Schutz, C. Maffeis, and A. P. Hills, "Childhood obesity and walking: guidelines and challenges," *International Journal of Pediatric Obesity,* vol. 6, pp. 332-341, 2011.
[23] S. P. Shultz, E. D'Hondt, P. W. Fink, M. Lenoir, and A. P. Hills, "The effects of pediatric obesity on dynamic joint malalignment during gait," *Clinical Biomechanics,* vol. 29, pp. 835-838, 2014.
[24] D. L. Gushue, J. Houck, and A. L. Lerner, "Effects of Childhood Obesity on Three-Dimensional Knee Joint Biomechanics During Walking," *Journal of Pediatric Orthopaedics,* vol. 25, pp. 763-768, 2005.
[25] P. G. Kopelman, I. D. Caterson, and W. H. Dietz, *Clinical Obesity in Adults and Children*: Wiley, 2009.
[26] R. C. Whitaker , J. A. Wright , M. S. Pepe , K. D. Seidel , and W. H. Dietz "Predicting Obesity in Young Adulthood from Childhood and Parental Obesity," *New England Journal of Medicine,* vol. 337, pp. 869-873, 1997.
[27] S. Harms, R. Larson, A. E. Sahmoun, and J. R. Beal, "Obesity increases the likelihood of total joint replacement surgery among younger adults," *International Orthopaedics,* vol. 31, pp. 23-26, 2007.
[28] M. W. Whittle, "Clinical gait analysis: A review," *Human Movement Science,* vol. 15, pp. 369-387, 1996.
[29] T. A. L. Wren, G. E. Gorton, S. Õunpuu, and C. A. Tucker, "Efficacy of clinical gait analysis: A systematic review," *Gait & Posture,* vol. 34, pp. 149-153, 2011.
[30] T. A. L. Wren, N. Y. Otsuka, R. E. Bowen, A. A. Scaduto, L. S. Chan, M. Sheng, R. Hara, and R. M. Kay, "Influence of gait analysis on decision-making for lower extremity orthopaedic surgery: Baseline data from a randomized controlled trial," *Gait & Posture,* vol. 34, pp. 364-369, 2011.
[31] E. H. Lee, J. C. H. Goh, and K. Bose, "Value of gait analysis in the assessment of surgery in cerebral palsy," *Archives of Physical Medicine and Rehabilitation,* vol. 73, pp. 642-646, 1992.
[32] V. Cimolin and M. Galli, "Summary measures for clinical gait analysis: A literature review," *Gait & Posture,* vol. 39, pp. 1005-1010, 2014.
[33] R. M. Ferrarin M., Bacchini M., Casiraghi A., Castagna A., Pizzi A., Montesano A. , "Does gait analysis change clinical decision-making in poststroke patients? Results from a pragmatic prospective observational study," *European Journal of Physical and Rehabilitation Medicine,* vol. 51, pp. 171-184, 2015.
[34] R. Baker, "Gait analysis methods in rehabilitation," *Journal of NeuroEngineering and Rehabilitation,* vol. 3, pp. 4-4, 2006.
[35] Y. Ehara, H. Fujimoto, S. Miyazaki, S. Tanaka, and S. Yamamoto, "Comparison of the performance of 3D camera systems," *Gait & Posture,* vol. 3, pp. 166-169, 1995.
[36] Y. Ehara, H. Fujimoto, S. Miyazaki, M. Mochimaru, S. Tanaka, and S. Yamamoto, "Comparison of the performance of 3D camera systems II," *Gait & Posture,* vol. 5, pp. 251-255, 1997.
[37] M. P. Kadaba, H. K. Ramakrishnan, and M. E. Wootten, "Measurement of lower extremity kinematics during level walking," *Journal of Orthopaedic Research,* vol. 8, pp. 383-392, 1990.

[38] E. Growney, D. Meglan, M. Johnson, T. Cahalan, and K.-N. An, "Repeated measures of adult normal walking using a video tracking system1Presented in part at the 1st Annual North American Clinical Gait Laboratory Conference, April 7–9 1994, Portland, OR, USA.1," *Gait & Posture,* vol. 6, pp. 147-162, 1997.
[39] T. Chau, S. Young, and S. Redekop, "Managing variability in the summary and comparison of gait data," *Journal of NeuroEngineering and Rehabilitation,* vol. 2, pp. 22-22, 2005.
[40] N. Stergiou, *Innovative Analyses of Human Movement*: Human Kinetics, 2004.
[41] G. Atkinson and A. M. Nevill, "Statistical Methods For Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine," *Sports Medicine,* vol. 26, pp. 217-238, 1998.
[42] A. Bruton, J. H. Conway, and S. T. Holgate, "Reliability: What is it, and how is it measured?," *Physiotherapy,* vol. 86, pp. 94-99, 2000.
[43] P. W. Stratford and C. H. Goldsmith, "Use of the Standard Error as a Reliability Index of Interest: An Applied Example Using Elbow Flexor Strength Data," *Physical Therapy,* vol. 77, pp. 745-750, 1997.
[44] N. E. Mayo, S. Wood-Dauphinee, R. Côté, L. Durcan, and J. Carlton, "Activity, participation, and quality of life 6 months poststroke," *Archives of Physical Medicine and Rehabilitation,* vol. 83, pp. 1035-1042, 2002.
[45] H. Y. Jang, Y. L. Kim, and S. M. Lee, "Perception and use of balance measures for stroke patients among physical therapists in South Korea," *Journal of Physical Therapy Science,* vol. 29, pp. 255-260, 2017.
[46] D. Meldrum, C. Shouldice, R. Conroy, K. Jones, and M. Forward, "Test–retest reliability of three dimensional gait analysis: Including a novel approach to visualising agreement of gait cycle waveforms with Bland and Altman plots," *Gait & Posture,* vol. 39, pp. 265-271, 2014.
[47] M. B. Pohl, C. Lloyd, and R. Ferber, "Can the reliability of three-dimensional running kinematics be improved using functional joint methodology?," *Gait & Posture,* vol. 32, pp. 559-563, 2010.
[48] B. Horsak, B. Pobatschnig, A. Baca, S. Greber-Platzer, A. Kreissl, S. Nehrer, B. Wondrasch, R. Crevenna, M. Keilani, and A. Kranzl, "Within-assessor reliability and minimal detectable change of gait kinematics in a young obese demographic," *Gait & Posture,* vol. 54, pp. 112-118, 2017.
[49] K. Kromeyer-Hauschild, M. Wabitsch, D. Kunze, F. Geller, H. C. Geiß, V. Hesse, A. von Hippel, U. Jaeger, D. Johnsen, W. Korte, K. Menner, G. Müller, J. M. Müller, A. Niemann-Pilatus, T. Remer, F. Schaefer, H.-U. Wittchen, S. Zabransky, K. Zellner, A. Ziegler, and J. Hebebrand, "Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben," *Monatsschrift Kinderheilkunde,* vol. 149, pp. 807-818, August 01 2001.
[50] S. E. Barlow, "Expert Committee Recommendations Regarding the Prevention, Assessment, and Treatment of Child and Adolescent Overweight and Obesity: Summary Report," *Pediatrics,* vol. 120, pp. S164-S192, 2007.
[51] M. Sangeux and J. Polak. Matlab implementation of the multiple functional depth with outliers detection. [Online].
[52] M. Sangeux and J. Polak. Matlab script to process the Excel spreadsheet in supplementary material and find the most representative stride. Now with outliers detection and output. [Online].
[53] M. Sangeux and J. Polak. Excel spreadsheet describing implementation and comparison of methods [Online].

[54] S. R. Simon, "Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems," *Journal of Biomechanics,* vol. 37, pp. 1869-1880, 2004.
[55] J. L. McGinley, R. Baker, R. Wolfe, and M. E. Morris, "The reliability of three-dimensional kinematic gait measurements: A systematic review," *Gait & Posture,* vol. 29, pp. 360-369, 2009.
[56] D. J. Rutherford, C. L. Hubley-Kozey, and W. D. Stanish, "The neuromuscular demands of altering foot progression angle during gait in asymptomatic individuals and those with knee osteoarthritis," *Osteoarthritis and Cartilage,* vol. 18, pp. 654-661, 2010.
[57] M. Simic, T. V. Wrigley, R. S. Hinman, M. A. Hunt, and K. L. Bennell, "Altering foot progression angle in people with medial knee osteoarthritis: the effects of varying toe-in and toe-out angles are mediated by pain and malalignment," *Osteoarthritis and Cartilage,* vol. 21, pp. 1272-1280, 2013.
[58] D. J. Rutherford, C. L. Hubley-Kozey, K. J. Deluzio, W. D. Stanish, and M. Dunbar, "Foot progression angle and the knee adduction moment: a cross-sectional investigation in knee osteoarthritis," *Osteoarthritis and Cartilage,* vol. 16, pp. 883-889, 2008.
[59] U. Della Croce, A. Leardini, L. Chiari, and A. Cappozzo, "Human movement analysis using stereophotogrammetry," *Gait & Posture,* vol. 21, pp. 226-237, 2005.
[60] G. Rab, K. Petuskey, and A. Bagley, "A method for determination of upper extremity kinematics," *Gait & Posture,* vol. 15, pp. 113-119, 2002.

List of Figures

List of Tables

Appendix

Table 7 The RMSD values averaged across the planes. ***AVG:*** *comparison between test and retest of averaged data sets across five trials.* ***MRT:*** *comparison between test and retest of two most representative data sets for each test session.* ***AVG-MRT1:*** *comparison between average curve across five trials and most representative trial for the test.* ***AVG-MRT2:*** *comparison between average curve across five trials and most representative trial for a retest. Data include mean ± SD.*

Mean (SD)	**F**	**S**	**T**
AVG	2,2 (1,04)	3,5 (2,41)	3,86 (2,41)
MRT	2,33 (0,97)	3,73 (2,09)	3,91 (2,2)
AVG-MRT 1	0,58 (0,2)	1,13 (0,56)	1,28 (0,4)
AVG-MRT 2	0,65 (0,19)	1,25 (0,72)	1,34 (0,66)

Table 8 The SEM values of kinematic parameters averaged across planes for each plane separately. MIN, MAX and ROM represent minimal and maximal values in the data set (curve amplitude) and range of motion (the difference between MIN and MAX), respectively. For each discrete parameter (MIN, MAX and ROM) comparison between test and retest was performed for averaged curves and most representative separately. Data include mean ± SD.

	AVG			**MRT**		
	MIN	MAX	ROM	MIN	MAX	ROM
F	1,48 (0,56)	1,95 (0,44)	1,92 (0,11)	1,5 (0,62)	1,93 (0,22)	1,6 (0,26)
S	3,12 (0,82)	2,91 (0,35)	2,06 (1,33)	3,17 (0,92)	2,59 (0,23)	2,23 (1,07)
T	3,28 (0,8)	3,57 (0,92)	2,07 (0,5)	2,85 (0,98)	3,06 (0,69)	2,56 (0,61)

Table 9 The SEM values of kinematic parameters averaged across planes for each plane separately. MIN, MAX and ROM represent minimal and maximal values in the data set (curve amplitude) and range of motion (the difference between MIN and MAX), respectively. For each discrete parameter (MIN, MAX and ROM) comparison between average curve across five trials and the most representative trial was performed for test and retest separately. Data include mean ± SD.

	AVG-MRT1			**AVG-MRT2**		
	MIN	MAX	ROM	MIN	MAX	ROM
F	0,36 (0,11)	0,51 (0,32)	0,53 (0,23)	0,38 (0,08)	0,57 (0,19)	0,64 (0,2)
S	0,97 (0,62)	0,68 (0,21)	1,01 (0,66)	0,93 (0,61)	1,1 (0,42)	0,95 (0,25)
T	0,92 (0,44)	0,99 (0,49)	1,35 (0,51)	1,68 (1,4)	1,55 (1,29)	1,38 (0,63)

Abbreviations

3DGA	Three-dimensional gait analysis
AVG	Average across five trials
BMI	Body Mass Index
GRF	Ground reaction force
FPA	Foot progression angle
ICC	Interclass correlation coefficient
LFM	Linear Fit Method
MAX	Maximal value in kinematic data set
MIN	Minimal value in kinematic data set
MRT	Most representative trial
RMSD	Root mean square deviation
ROM	Range of motion
SEM	Standard error of measurement
sEMG	Surface electromyography