A segmentable normative database for gait analysis based on the Helen Hayes markerset

Paula Flórez Herrera[1,*8] & Pablo Callejo[1]
José Luis Luque[*2], Manuel Miranda[*3], Lluis Guirao[*4], Ernesto Lucas[*5], Rocío Unzurrunzaga[*5], Carme Unyó Sallent[*5], Margarita Sauné Castillo[*5], Jesús García[*6] & Antonio Donaire[*7]

Abstract— **This study encompasses the planning, data collection, processing and creation of the first Spanish normative database for gait analysis. The current document summarizes the experimental procedure followed by the data collection personnel with each of the subjects, followed by the protocol used for data processing and the final merger of the finished normative dataset.**

The entirety of the data was collected by the institutions listed below (hospitals or research centers located in Spain and users of STT applications), after a thorough training provided by STT Systems, and all the resulting captures were processed at the STT Systems headquarters. The initial goal was to obtain enough captures to be able to segment the database by gender, age group and/or footwear without compromising its robustness. Finally, the resulting database was embedded into the motion analysis applications involved, in order to make it available for the users.

I. INTRODUCTION

The importance of normative databases in the field of gait analysis goes beyond having a normal reference to check whether the data collected varies from the norm or not. Normative databases are tools that can help physicians understand the intricacies of human gait and the nature of ability and disability itself.

For as long as gait analysis has existed as a field, the use of a normative database as a contrasting element of the classic gait graphs has been a standard. Many databases for normal human gait reported in the literature are of free access, and these same databases are the ones commonly used by applications for gait analysis, the 3DMA products of STT Systems being one of them [1].

This study uses optical motion capture as the method of data collection, specifically, STT Systems' 3DMA products. With the widespread use of the 3DMA products for research in both the clinical and sports fields, came the understanding that the normative database currently in use for gait analysis within the products held a certain amount of bias due to a couple of factors.

First, there was a geographical bias due to the headquarters of STT Systems being in Spain (as well as a sizeable number of the users of 3DMA applications), while the normative database currently being used was compiled in its entirety in the United States using North American subjects. The second, the unavoidable bias of the normative data being collected with a different application for motion analysis. This last issue in particular led STT Systems to propose the creation of a new normative dataset for gait analysis using 3DMA applications.

In order to collect a complete database for normal gait with enough subjects to be used as a contrasting point for clinical data, STT Systems decided to team up with a group of hospitals and research facilities that could collect data from subjects from different parts of Spain and then proceed to merge it into a single database that could be available in 3DMA applications.

II. EXPERIMENTAL PROCEDURE

A. *Examination Preparation*

From the beginning of the project, it was crucial that the normative data was collected as homogeneously as possible throughout the different test subjects and research facilities involved in the process, in order to minimize both intra and inter-user variability and ensure less biased results for the final database. To participate in the project, the institutions were required to set aside a slot of time for the examiners to go through specialized training, provided by STT Systems, and then commit to collect and send the walking data of around 10 subjects following the experimental protocol. The final group of collaborators and their locations is listed in Table I.

TABLE I
DATA COLLECTION INSTITUTIONS

Institution	Location
Mutua Navarra[*2]	Pamplona, Spain
FREMAP Majadahonda[*3]	Madrid, Spain
Hospital de Mataró[*4]	Catalonia, Spain
MC Mutual[*5]	Barcelona, Spain
FREMAP Sevilla[*6]	Sevilla, Spain
BioGait[*7]	Granada, Spain
STT Systems[*8]	Gipuzkoa, Spain

This is an internal study and its results are not meant for publication or use outside of the company or its authorized clients.

[1]Motion Capture Division, STT Systems. www.stt-systems.com San Sebastian, Spain.

*Responsible for data collection for the project at their institution. Consult indexes in Table I.

The content of the training included the following key aspects:

- Current normative database used by the gait analysis protocols of 3DMA applications.
- Differences between normative databases currently available for free use.
- Content of the informed consent forms for the subjects to sign, as well as the patient questionnaire that the examiner must fill out with statistical data such as age, weight, height and gender of each subject.
- Specifications of the markerset that will be used for the data collection of all subjects (correct marker placement, common mistakes and how to spot them), as well as a hands-on training where the examiners practiced the marker placement and capture process before moving on to the test subjects (Fig. 1).
- Exclusion criteria, in order to obtain a normative dataset containing subjects with no known pathologies or surgical record that could alter their gait pattern.
- Instructions on the correct calibration of the capture volume before the tests (the data from the calibration was saved for each trial in the study).

Fig. 1. Pictures from the training provided by STT Systems to the research institutions collaborating with the project in October of 2017: Motion capture laboratory teams from FREMAP Majadahonda (top) and Hospital de Mataró (bottom).

B. Markerset

The protocol used for data collection was based on the original Helen Hayes markerset for lower-body kinematics analysis [2]. This protocol, from now on referred to as the STT Helen Hayes protocol (Fig. 2), has been included as an option for the optical capture of gait in the STT software (3DMA products) for the past 15 years.

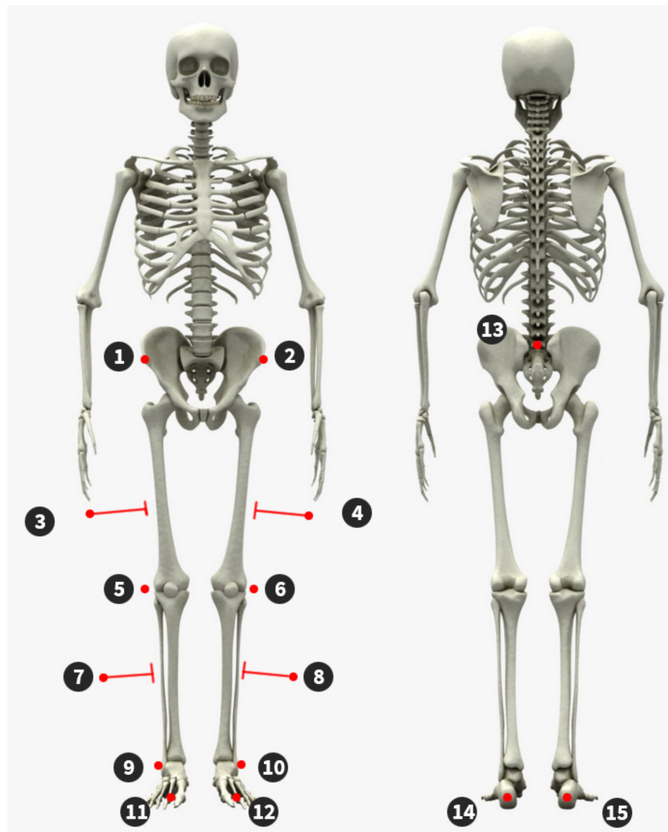

Fig. 2. Anatomical protocol used for the optical system marker placement.

The main difference between the original markerset and the modified one was the removal of the markers on the greater trochanters (GT). Since the calculation used to find the hip centers of rotation requires only the Anterior Superior Illiac Spine (ASIS) markers [3], the GT markers were subtracted from the final setup and a marker on each calcaneous prominence was added instead (heel markers). The resulting markerset, STT Helen Hayes, is listed in Table II.

The addition of the heel markers allows for the tridimensional tracking of the foot, which was not possible with the original markerset. One of the particularities of the normative database previously available in the 3DMA products was the fact that there was no normal curve for the ankle abduction/adduction. With the tracking of the heel in this protocol, there are enough tracking points in the shank and foot to obtain a curve for the movement of the ankle joint in the coronal plane.

TABLE II

MARKERS IN THE STT HELEN HAYES PROTOCOL

Marker name	Anatomical landmark
1. Right ASIS	Prominence of the right anterior superior illiac spine
2. Left ASIS	Prominence of the left anterior superior illiac spine
3. Right thigh	Lateral side of the right thigh (with a plastic extender)
4. Left thigh	Lateral side of the left thigh (with a plastic extender)
5. Right knee	Lateral epicondyle of the right femur
6. Left knee	Lateral epicondyle of the left femur
7. Right thigh	Lateral side of the right shank (with a plastic extender)
8. Left thigh	Lateral side of the left shank (with a plastic extender)
9. Right ankle	Right lateral malleolus
10. Left ankle	Left lateral malleolus
11. Right toes	Between the 2nd and 3rd metatarsals of the right foot
12. Left toes	Between the 2nd and 3rd metatarsals of the left foot
13. Sacrum	Over the S2 vertebrae
14. Right heel	Ball of the calcaneus of the right foot
15. Left heel	Ball of the calcaneus of the left foot

Before the walking trials, subjects were required to stand in anatomical position for a static capture as an initial calibration of the model. The remainder of the biomechanic calculations follow the original Helen Hayes model [2], and the resulting curves are shown in Table III.

TABLE III

JOINT ANGLES

Joint	Sagittal Plane	Coronal Plane	Transverse Plane
Pelvis	Anterior/Posterior Tilt	Superior/Inferior Obliquity	Internal/External Rotation
Hip	Flexion/Extension	Abduction/Adduction	Internal/External Rotation
Knee	Flexion/Extension	Varus/Valgus	Internal/External Rotation
Ankle	Flexion/Extension	Abduction/Adduction	Internal/External Rotation
Foot	-	-	Internal/External Rotation

C. *Testing criteria*

In order to ensure that the normative database was in fact, made entirely out of normal data, the following aspects were deemed exclusion criteria for the test:

- The subject presents bone deformities in lower limbs.
- The subject presents exaggerated muscle retractions, ligamentous laxity, or spasticity.
- The subject has endured one or more surgical interventions for the lower limbs or back in the past 5 years.
- The subject has a history of persistent injury (sports or otherwise) in the lower limbs.
- The subject is diagnosed or suspected of presenting any other pathology that the examiner deems could potentially affect their gait pattern.

After going through the initial examination to rule out the exclusion criteria, the subjects filled the necessary forms, then were measured, weighted and suited with the 15 markers required for the captures.

The subjects were instructed to walk normally, relaxed, and in a straight line along the capture area, making sure they had a few tries to settle to the feeling of the markers before the tests were recorded. Subjects were recorded for as long as possible, resulting in tests varying from 10 to up to 100 strides per capture.

III. DATA PROCESSING

A. *Acquisition*

After receiving the training provided by STT Systems, the research institutions proceeded to collect walking data from subjects following the specifications provided above for a period of 8 months.

The final tally for September of 2018 was of 73 captures from 59 test subjects (mean weight = 71.9 ± 17.2 kg, mean height = 171 ± 8 cm), adding up to a total of 3123 strides. Once STT received all the captures, the Motion Capture division proceeded with the examination and filtering of the data.

B. *Filtering*

The first step of processing for each of the captures received from the research institutions was the filtering of the data, this was done using the *stride selection dialog* available for users of 3DMA applications for gait and running (Fig. 3).

This dialog uses the automatic calculation of gait events provided by the software (foot initial contacts and toe-offs) to divide each joint curve (Table III) into strides (beginning and ending with the initial contact of each respective foot to the ground, describing the entirety of the gait cycle from 0 to 100%.

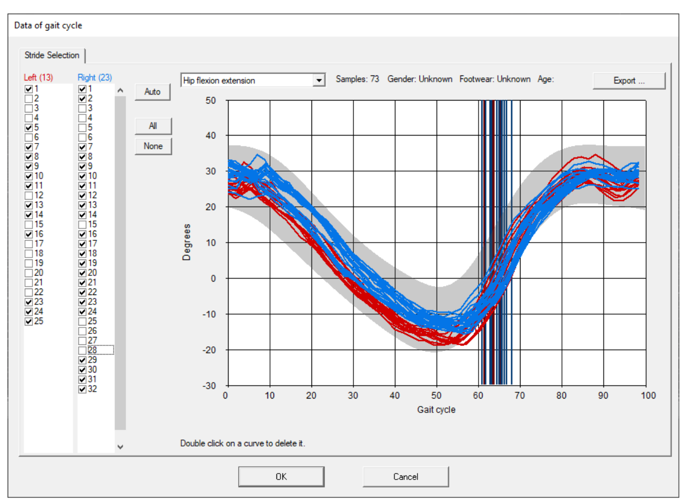

Fig. 3. Stride selection dialog from 3DMA applications used to filter out the *turning strides* or noise found in the captures. This example shows the Hip Flexion/Extension angles of a sample subject from the study.

The process of filtering was required for two specific cases: turning stride and noise. A turning stride happens when the subject makes a 180° turn at the end of the walkway in order to continue walking in the opposite direction. When this happens inside the capture volume, the gait events will be calculated during the turn and the data obtained from that stride will not correspond to a normal gait pattern.

On the other hand, a stride with noise could happen during the normal gait pattern but for any reason (marker occlusion, reflections in the capture volume, etc.) the curve shows one or several peaks that do not accurately represent the performed movement. All these strides were eliminated from the trials to avoid skewing the data.

The curves shown on Fig. 3 correspond to the Hip Flexion/Extension angle of one of the subjects included in the study. The plot window also shows the normative data for Hip Flexion/Extension in the background (grey shadow). The filtering of the captures was done for all the angles in Table III (by selecting them individually in the stride selection dialog from Fig. 3 and deselecting or ticking off the unwanted curves).

C. Data Export

Once each capture was checked and filtered, the data was exported to CSV from the same *stride selection dialog* in Fig. 3. Each capture yields 26 CSV files named accordingly to each joint angle (Table III) for right and left side of the body. Inside each CSV file, the data is structured in columns, one for each valid stride with 101 rows (from 0 to 100% inclusive in 1% intervals).

The resulting data was then imported to GNU Octave [4] in the form of matrices and the left and right sides were combined into 13 matrices total, each containing the total amount of strides of each trial as well as the statistical information from the trial (subject's age, gender, footwear and ID).

D. Segmentation

The groups for the statistical segmentation of the data are presented in Table IV. The data was segmented not only for statistical comparison of the different groups, but also to provide the option of normative database filtering for 3DMA users, in order to compare gait trials from patients with a database tailored for them.

The graphical comparison between different segments of the dataset presented in the results section was done in Octave, however, none of these pre-calculated sub-datasets were used directly in 3DMA applications. The software contains the entirety of the dataset by trial and the segmentation (if required) is done within the *normative dataset* dialog.

TABLE IV
SEGMENTATION OF THE FINAL DATASET

Variable	Segment	# of trials	# of strides
Gender	Male	38	1841
	Female	27	1096
	N/D*	8	186
Age group	18-29	21	863
	30-44	20	732
	45-61	16	754
	N/D*	16	774
Footwear	No shoes	33	1792
	Shoes	32	1145
	N/D*	8	186

*N/D: Subjects wished not to disclose the statistic or all optional statistical questions in the consent form.

E. Merge

The final merger of the resulting dataset yields 13 joint angles (Table III), each containing 101 values for mean of all the strides and 101 for standard deviation of all the strides (from 0 to 100% inclusive in 1% intervals). The graphical representation of these normative curves is presented in the following section.

IV. RESULTS

A. Pelvis

For the pelvic joint (Fig. 4), the angles are measured against a fixed set of coordinates (such as the subject calibration data or static trial), except for the rotation, which is measured against the gait direction of movement (the case for all subjects in this study) or the treadmill direction (if the subject were to walk on a treadmill).

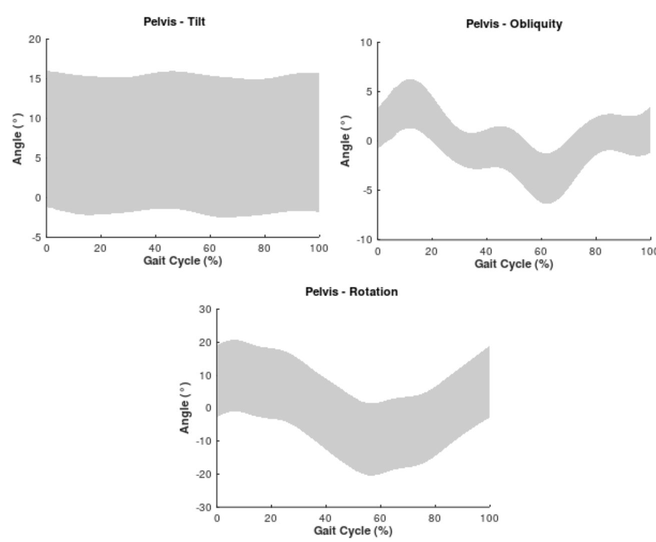

Fig. 4. Resulting normative curves for the pelvic joint. Pelvic tilt (anterior-posterior) in the sagittal plane (top left), pelvic obliquity (superior-inferior) in the coronal plane (top right) and pelvic rotation (internal-external) in the transverse plane (bottom).

B. Hips

The hip joint angles (Fig. 5) were calculated by comparing the pelvic orientation with the femur orientation in the hip center of rotation [3].

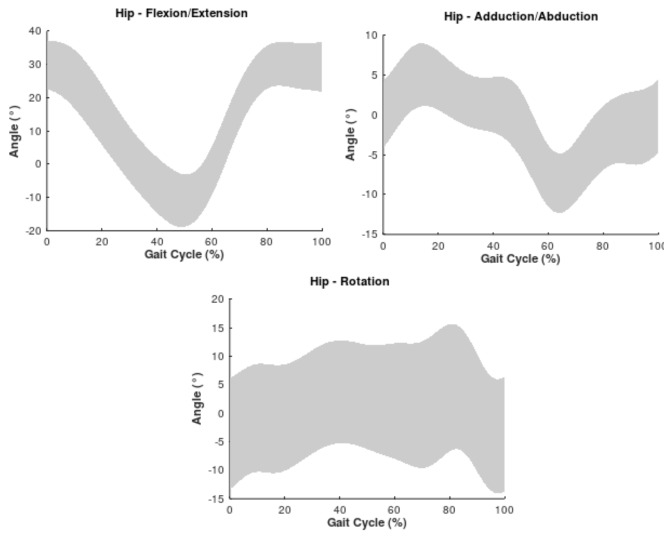

Fig. 5. Resulting normative curves for the hip joints. Hip flexion (flexion-extension) in the sagittal plane (top left), hip abduction (adduction-abduction) in the coronal plane (top right) and hip rotation (internal-external) in the transverse plane (bottom).

C. Knees

The knee joint angles (Fig. 6) were calculated by comparing the femoral orientation with the tibial orientation in the knee center of rotation [2].

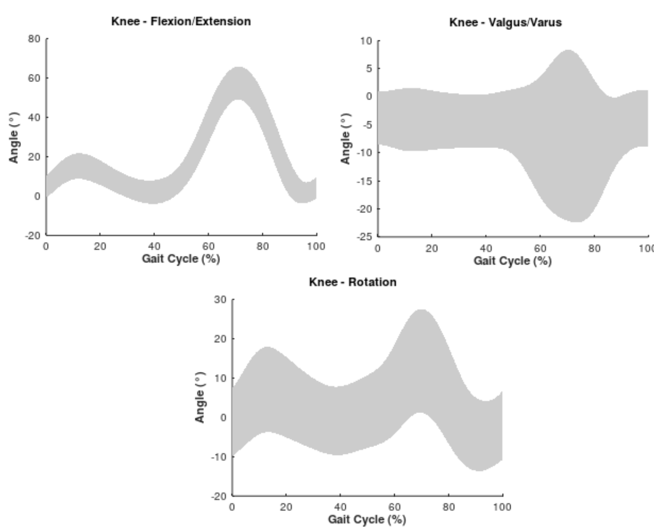

Fig. 6. Resulting normative curves for the knee joints. Knee flexion (flexion-extension) in the sagittal plane (top left), knee abduction (valgus-varus) in the coronal plane (top right) and knee rotation (internal-external) in the transverse plane (bottom).

D. Ankles

The ankle joint angles (Fig. 7) were calculated by comparing the tibial orientation with the foot orientation in the ankle center of rotation [2].

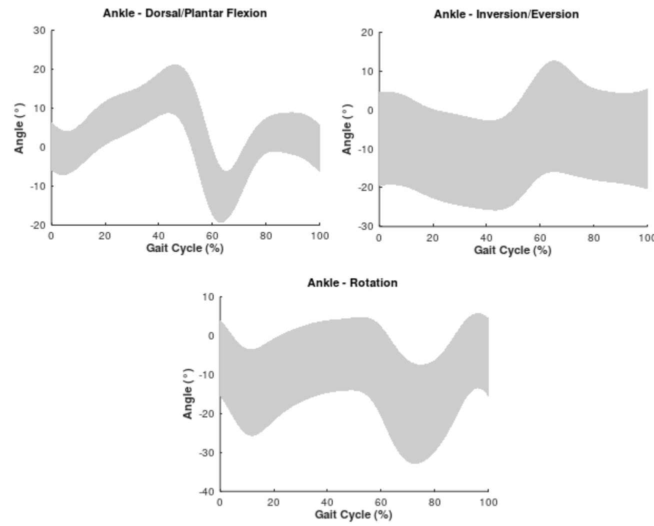

Fig. 7. Resulting normative curves for the ankle joints. Ankle flexion (dorsal-plantarflexion) in the sagittal plane (top left), ankle abduction (inversion-eversion) in the coronal plane (top right) and ankle rotation (internal-external) in the transverse plane (bottom).

E. Feet

The foot direction (Fig. 8) was calculated by comparing the foot orientation with the pelvic orientation in the pelvic center of mass. It describes a global rotation of the leg throughout the gait cycle.

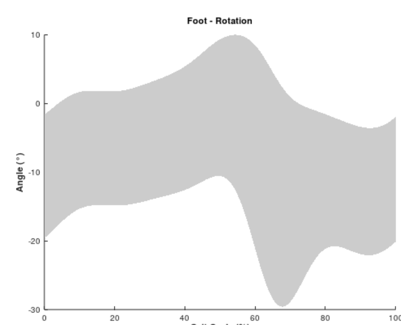

Fig. 8. Resulting normative curves for the foot rotation (internal-external) in the transverse plane.

V. CONCLUSIONS

Advanced measurement tools such as optical tracking systems have the potential of providing very complex data to their users, to the point where a normative database becomes an essential guide for the interpretation and analysis of said data.

The main goal of this study was to create a normative database large and diverse enough to able to be segmented. The importance of the segmentation of the database comes from the observation and analysis of other gait analysis databases currently available for free use, and the differences between them. Given that some of these databases do not provide the statistical information of the subjects used to collect the data, it is impossible to determine if the difference between them is rooted on some kind of systematic bias or if it is merely a product of chance.

STT Systems considered of the utmost importance to obtain a normative database that could allow the user to rule out statistical bias as much as possible, as well as guaranteeing that the entirety of the database was collected using the same motion captor system and biomechanical protocol their customers use, diminishing the possible systematic differences from the normative that could have been present before.

As it is shown in annex I and II, some segmentations will yield more significant difference between them than others. In the case of the first annex, not many general differences can be observed between the gait curves of men and women, but a few specific details can be identified as the telltale of the segments, for example, the noticeable pattern of increased hip internal rotation before toe-off in women (Fig. 9: Hip Rotation, upper red line) that indicates a sustained different behavior in the hip of the load bearing leg while the body weight is being sustained by it, visibly seen as swaying the hips during the gait cycle. Some other segmentations will show many more differences, such as the case of annex II (Fig. 10), that compares the youngest and oldest subjects of the study. These differences could indicate a change of the gait pattern with aging, just as much as they could be biased by the size of the samples (2 person groups for each segment).

Finally, the creation of this normative database opened a door for the collection of normative data as a tool for biomechanical analysis in STT, and working with professionals who use the applications on a daily basis has helped the development team at STT Systems to evolve the software and its newest *normative selector* feature to fit the needs of the application users.

VI. FUTURE WORK

Even with a large database such as the one resulting from in this study, a triple filter like this one can diminish the overall number of trials used for the final comparison to the gait data. STT plans to continue collecting normative data in order to increase the final tally of trials in the database and contribute to its solidness. Ideally, a database with 100 trials would be large enough to endure multiple layers of segmentation and remain robust.

Another feature currently in the works at STT Systems is a tool for the users to create their own normative databases. The goal would be to facilitate not only the creation of a local normal gait reference for the users interested in doing so, but the option to create multiple databases for reference besides a normal one, for example, databases for specific gait pathologies or even databases for movements beyond gait.

REFERENCES

[1] STT Systems, San Sebastian, Spain. Website: www.stt-systems.com.

[2] Kadaba, M. P., Ramakrishnan, H. K., & Wootten, M. E. (1990). Measurement of lower extremity kinematics during level walking. Journal of Orthopaedic Research,8(3), 383-392. doi:10.1002/jor.1100080310

[3] Bell, A. L., Brand, R. A., & Pedersen, D. R. (1989). Prediction of hip joint centre location from external landmarks. Human Movement Science,8(1), 3-16. doi:10.1016/0167-9457(89)90020-1

[4] GNU Octave, version 4.4.1 Copyright (C) 2018 John W. Eaton and others.