Journal of Diabetes Science and Technology
Volume 4, Issue 5, September 2010
© Diabetes Technology Society

SYMPOSIUM

A New Classification of Diabetic Gait Pattern Based on Cluster Analysis of Biomechanical Data

Zimi Sawacha, Ph.D.,[1] Gabriella Guarneri, M.D.,[2] Angelo Avogaro, M.D., Ph.D.,[2] and Claudio Cobelli, Ph.D.[1]

Abstract

Background:

The diabetic foot, one of the most serious complications of diabetes mellitus and a major risk factor for plantar ulceration, is determined mainly by peripheral neuropathy. Neuropathic patients exhibit decreased stability while standing as well as during dynamic conditions. A new methodology for diabetic gait pattern classification based on cluster analysis has been proposed that aims to identify groups of subjects with similar patterns of gait and verify if three-dimensional gait data are able to distinguish diabetic gait patterns from one of the control subjects.

Method:

The gait of 20 nondiabetic individuals and 46 diabetes patients with and without peripheral neuropathy was analyzed [mean age 59.0 (2.9) and 61.1(4.4) years, mean body mass index (BMI) 24.0 (2.8), and 26.3 (2.0)]. *K*-means cluster analysis was applied to classify the subjects' gait patterns through the analysis of their ground reaction forces, joints and segments (trunk, hip, knee, ankle) angles, and moments.

Results:

Cluster analysis classification led to definition of four well-separated clusters: one aggregating just neuropathic subjects, one aggregating both neuropathics and non-neuropathics, one including only diabetes patients, and one including either controls or diabetic and neuropathic subjects.

Conclusions:

Cluster analysis was useful in grouping subjects with similar gait patterns and provided evidence that there were subgroups that might otherwise not be observed if a group ensemble was presented for any specific variable. In particular, we observed the presence of neuropathic subjects with a gait similar to the controls and diabetes patients with a long disease duration with a gait as altered as the neuropathic one.

J Diabetes Sci Technol 2010;4(5):1127-1138

Author Affiliations: [1]Department of Information Engineering, University of Padova, Padova, Italy; and [2]Department of Clinical Medicine and Metabolic Disease, University Polyclinic, Padova, Italy

Abbreviations: (ANOVA) analysis of variance, (BMI) body mass index, (CS) control subjects, (DPN) diabetes patients with diabetic neuropathy, (GRF) ground reaction force, (NoDPN) diabetes patients without diabetic neuropathy

Keywords: classification, cluster, diabetes, kinematics, kinetics, three dimensional

Corresponding Author: Claudio Cobelli, Ph.D., Department of Information Engineering, University of Padova, Via Gradenigo 6B, Padova 35131, Italy; email address cobelli@dei.unipd.it

Introduction

The World Health Organization warned that, in 2000, as many as 33 million Europeans suffered from diabetes, and approximately 15% of these patients will likely develop foot ulcers and approximately 15% to 20% will face lower-extremity amputation.[1] The diabetic foot, one of the most serious complications of diabetes mellitus and a major risk factor for plantar ulceration, is determined mainly by peripheral neuropathy, foot trauma, foot deformity, increased foot pressures, and callus.[1,2] Distal symmetric sensorimotor polyneuropathy is primarily confined to the axons of small- and large-fiber sensory afferents. It can be defined as the presence of symptoms and/or signs of peripheral nerve dysfunction in people with diabetes, after exclusion of other causes.[3] The result is a "stocking feet" pattern of sensory loss that begins in the toes and progresses proximally.[3] Peripheral neuropathy patients exhibit decreased stability while standing[4–6] as well as during dynamic conditions.[7] Several authors[8–13] found gait pattern alteration in the sagittal plane in diabetes patients with diabetic neuropathy (DPN) and in diabetes patients without diabetic neuropathy (NoDPN).

Nevertheless, alterations of the proprioceptive system, like DPN, could affect walking not only because there is reduced peripheral sensory information available to local peripheral systems but also because this reduced information imposes a burden on the higher cortical centers involved in the processing of sensory information.

Several hypotheses suggest that the integrity of higher cortical/central factors is of significance for posture and gait control.[12] Authors[14,15] have identified changes in some gait parameters that appear to be specific in diabetes: shorter stride length, reduced walking speed, and altered lower limb and trunk mobility. These observations highlight the need for an objective and quantitative measurement of the relevant aspects of the gait of diabetes patients. Such data should ultimately lead to a better understanding of interrelated foot/limb function, foot management, and both kinetic and kinematics elements during gait.

Despite the importance of diabetes in gait alterations, previous research examining gait in diabetic subjects adopted ensemble averages as a standard method of describing the kinematics and kinetic patterns observed. A problem inherent with the use of ensemble averaging is that significant variations in both kinematics and kinetics patterns in each population of subjects (for instance, DPN, control subjects (CS), and NoDPN) may not be observed. Thus an ensemble average may not provide an accurate representation of the gait seen across individuals. Cluster analysis is a technique developed to identify natural groupings that may exist in a population of interest.[16] This technique places individuals into a cluster that contains other individuals with similar characteristics. Similarities and dissimilarities are determined by the analysis of descriptive variables for each individual. Most gait studies involving DPN and NoDPN subjects have been performed grouping population by age, body mass index (BMI), and diagnosis of neuropathy parameters. To our knowledge, no studies have been done on diabetic subjects stratified for biomechanical parameters; meanwhile, this has been performed in healthy young adults and elderly.[17–19] Here we describe DPN biomechanical impairment during walking, grouping population stratified by their gait patterns. Furthermore, cluster analysis has the advantage that several parameters can be taken into account at the same time rather than a single one for each individual.[18] Thus, in this study, it was also employed to identify if there were some gait parameters able to distinguish the gait patterns of diabetic or neuropathic subjects from those of controls. In this case, cluster analysis can be used to identify from the parameters determined with gait analysis those that best describe gaits of diabetic subjects. Considering that cluster analysis divides data into meaningful or useful groups (clusters), if meaningful clusters are the goal, then the resulting clusters should capture the "natural" structure of the data. As opposed to classical data interpretation, relying on human *a priori* knowledge and experience and therefore being exposed to biased results and to missed valuable information, cluster analysis could be employed as a data mining technique. We define data mining to be the discovery of useful, but nonobvious, information or patterns in large collections of data. Data mining has been shown to provide a more complete and unbiased picture of the phenomenon under investigation.[20,21]

Therefore the aims of this study were to (1) verify if a standard method of describing the kinematics and kinetic patterns by means of ensemble averages could be adopted efficiently in the diabetic population, (2) verify if three-dimensional gait data would be able to distinguish

the gait patterns of DPN and NoDPN subjects from those of CS, and (3) identify both clinical and biomechanical characteristics of the natural groupings observed in the population of interest.

Materials and Methods

Patient Recruitment

Subjects were consecutively recruited among the patients attending the outpatient clinic of the Department of Metabolic Disease of the University of Padova, Italy. Inclusion criteria were type 1 and 2 diabetic subjects with walking ability and no history of ulcers, neurological disorders (apart from DPN), orthopedic problems, or lower limb surgery. The CS were recruited among hospital personnel. On the basis of these criteria, 66 patients were evaluated; 20 nondiabetic individuals [CS, mean age 59.0 (2.9), mean BMI 24.0 (2.8)], 20 NoDPN [mean age 63.8 (5.4), mean BMI 26.3 (2.0)], and 26 DPN [mean age 63.2 (6.0), mean BMI 25.6 (2.9)] were selected. All subjects gave written informed consent. The protocol was approved by the local ethics committee. The sample size was calculated with Altman Nomogram.[22]

Clinical Assessment and Postural Examination

The feet were checked for skin lesions, bone deformities, ulcerations, signs of infection, and previous amputations. Height (m) and weight (wearing only slips, without shoes) were recorded and BMI (kg/m^2) calculated.

The neurological evaluation included the assessment of symptoms and signs compatible with peripheral nerve dysfunction. The Michigan Neuropathy Screening Instrument questionnaire, which evaluates motor and sensory symptoms (subjects were classified as pathologic if 3 positive scores out of 15 were found), was filled.[23] The physical examination consisted of (1) patellar and ankle reflexes, with the patient in the sitting position; (2) assessment of muscle strength by ability to walk on heels, bilateral dorsiflexion–plantar flexion of the feet, flexion–extension of legs, abduction–adduction of both forearms and fingers, all against resistance; (3) sensory testing carried out on the index finger and on the hallux (pinprick with a disposable 25/7 mm needle), touch (10 g Semmens Weinstein monofilament, pathologic if no response on 3 out of 10 sites), and vibration perception threshold (128 MHz tuning fork and biothesiometer, pathologic if >25 V); (4) pain sensitivity; (5) electroneurophysiological study; and (6) Winsor index (ankle-to-brachial index). The cardiovascular autonomic tests were also performed, namely, deep breathing test, lean-to-standing test, Valsalva maneuver, and orthostatic change in blood pressure. If two or more tests were abnormal, the patient was considered to have autonomic neuropathy.

Hemoglobin A1c values in the preceding 10 years were collected. Each patient had at least an ophthalmologic examination, a urinary albumin-to-creatinine ratio (0–30 mg/g normal, 30–300 mg/g microalbuminuria, >300 mg/g macroalbuminuria), a carotid artery Doppler examination, and an electrocardiogram in the preceding three months.

Subjects underwent foot morphological examination to assess type of foot (cavus, planus, normal), foot deformity evaluation (hallux valgus/normal/rigitus, claw and hammer toes, limitation of dorsiflexion of the great toe fifth abducted/adducted toe, overlapping toes), and pre–post surgery ulcers lesions (bunions, callosity, scars, soft corn). Hip–knee–ankle joint mobility was also assessed. The magnitude of sagittal curvature was measured (considering the pelvic structure on the sagittal plane, both a forward or a backward displacement of the point tangential to the scrum–coccyx plane was estimated), and the pelvic cingulum (pelvic translation, pelvic rotation, pelvic bascule), cingulum scapular rotation (rotation of the clavicular acromion with respect to the contralateral side), cingulum scapular flexion (fall of the clavicular acromion with respect to the contralateral side), leg length asymmetry (a tape is placed on the umbilicus and stretched to each medial malleolus and a measure recorded, and the differences between the right and left leg measurements are then considered) superior of 5 mm, heel position, and plantar foot arch during bipedal load were evaluated.[15,24,25]

Motion Analysis Methodology

Gait analysis was performed using a 120–160 Hz six-camera stereophotogrammetric system (BTS S.r.l, Padova) and two force plates (Bertec Corporation). The signals coming from all systems were synchronized. Fifty-two reflective markers were placed on the subjects in correspondence of either anatomical landmarks or technical clusters of trunk, thigh, shank, foot, and pelvis (six clusters, each of them formed by four markers).[15] The protocol is based on the CAST technique[26] and represent a modified existing version.[15,27] Almost each anatomical landmark was calibrated using a static

acquisition without the aid of a pointer.[15,26] The following anatomical landmarks were considered for direct marker placement:

- Trunk—right and left acromions, spinous process of seventh cervical vertebrae (C7), spinous process of fifth lumbar vertebrae (L5).
- Foot—right and left calcaneus, right and left first metatarsal head, right and left second metatarsal head, right and left fifth metatarsal head.

The following anatomical landmarks were considered for direct marker placement and were calibrated with respect to a local cluster of marker by means of a static acquisition:

- Thigh—right and left lateral and medial epicondyle.
- Shank—right and left tibial tuberosity, right and left head of the fibula, right and left lateral and medial malleolus.

Four extra markers were applied on the thigh, pelvis, and shank in order to create the clusters that were used to calibrate the position of each segment anatomical landmarks.[15,26]

The following anatomical landmarks were calibrated with the aid of a pointer:[15,26]

- Pelvis—right and left anterior superior iliac spine, right and left posterior superior iliac spine.
- Thigh—right and left greater trochanter.
- Shank—right and left tibial tuberosity, right and left head of the fibula, right and left lateral malleolus, right and left medial malleolus.

The center of the femoral head was assumed to coincide with the center of the acetabulum. The latter was reconstructed by means of a functional method proposed by Cappozzo and colleagues.[26]

Anatomical reference frames for the body segments were defined according to previous work.[15,27] Standard coordinate systems[28] were adopted for each joint, which entails defining flexion–extension as the relative rotation about mediolateral axis of the proximal segment, internal–external rotation as the relative rotation about the vertical axis of the distal segment, and abduction–adduction as the relative rotation about a "floating" axis orthogonal to these two at each collected sample. The joint angles considered for the kinematics analysis were the flexion–extension, abduction–adduction, and internal–external rotation of trunk, pelvis, hip, and ankle. In particular, when considering the ankle joint, these three rotations were referred to respectively as dorsiflexion–plantar flexion, inversion–eversion, and internal–external rotation. With respect to the knee in the report of joint rotations, only the flexion–extension angle was reported; abduction–adduction and internal–external rotation were not considered although the model accounts for their values, because these joint rotations were not shown to be feasible when reconstructed through markers placed directly on the skin.[15,29] In the postprocessing of the kinetics parameters, the flexion–extension, abduction–adduction, and internal–external rotation moments of trunk, hip, knee, and ankle were considered together with the mediolateral, vertical, and anteroposterior forces.[15]

During gait analysis, the patients walked at a self-selected speed along a runway. Subjects were asked to walk along the walkway so that the target foot would naturally land on the force plates. Three walking trials, with three right and three left foot contacts on the force plate, were conducted.

Velocity, stride, and step parameters (namely, gait velocity, stride length, stride period, and step period) were calculated together with all angular displacements and internal joint moments.[15] A static acquisition was also performed.[14]

Kinematic and kinetic parameters were estimated and compared with those obtained with the data of a control population (mean age of 59 ± 2.9 years, mean BMI of 24 ± 2.8) by means of *k*-means cluster analysis.[30] The function *kmeans* (Matlab software, R2008b) was used, and the standard euclidean distance was chosen in forming the clusters. *kmeans* uses an iterative algorithm that minimizes the sum of distances from each object to its cluster centroid over all clusters. This algorithm moves objects between clusters until the sum cannot be decreased further. In this context, the optimal number of cluster was identified with the following procedure:

1. The first number of cluster is guessed (in this case, the first number of cluster chosen was equal to 2).

2. The optional "display" parameter (Mathlab software) is used to print information about each iteration, and the total sum of distances is evaluated. The latter may vary at each iteration as *kmeans* reassigns points between clusters and recomputes cluster centroids.

3. A silhouette plot for each solution is generated in order to evaluate which solution generates clusters that are better separated than previous solutions. The silhouette value for each point is a measure of how similar that point is to points in its own cluster compared to points in other clusters. This measure ranges from +1, indicating points that are very distant from neighboring clusters, through 0, indicating points that are not distinctly in one cluster or another, to -1, indicating points that are probably assigned to the wrong cluster. It is defined as

$$S(i) = (\min(b(i,:),2) - a(i))/\max(a(i),\min(b(i,:),2)),$$

where $a(i)$ is the average distance from the ith point to the other points in its cluster, and $b(i,k)$ is the average distance from the ith point to points in another cluster k.

4. In order to overcome the problem that *kmeans* solution often depends on the starting points, the optional "replicates" parameter (Mathlab software) is used. Thus the number of "replicates" is indicated, and the "display" parameter (Matlab software, R2008b) to print out the final sum of distances for each of the solutions is used. So far, the final solution that *kmeans* returns is the one with the lowest total sum of distances over all replicates.

5. The number of cluster is increased at each solution until an empty cluster is created.

6. The solution that generates clusters that are better separated than previous solutions is chosen.

The following variables were considered: joint moments (trunk, hip, knee, ankle), joint angles (trunk, hip, knee, ankle), and ground reaction force (GRF) components (anterior–posterior, mediolateral, vertical) for their values in NoDPN, DPN and CS. The knee internal–external and abduction–adduction angles were not considered as in previous studies.[15] So far, the clustering *kmeans* algorithm was applied to 1675 curves (each curve represents 100 samples, which are the variable's value in each percentage of the gait cycle or the stance phase of gait). Through visual inspection of each silhouette plot, the optimal solution was defined for each variable. Afterward, the clusters generated by the variables in which *kmeans* reaches an optimal solution were chosen and analyzed; the mean, range, and standard deviation of kinematics, kinetics, and time and space parameters of each cluster were computed, together with the clinical outcome.

One-way analysis of variance (ANOVA) (Matlab software, R2008b) was performed for joint rotation angles, moments, and forces in order to identify significant differences ($p < .05$) among each cluster's set of variables. Hence, one-way ANOVA[14] was computed between paired groups of subjects: DPN versus NoDPN, DPN versus CS, NoDPN versus CS, and between subjects in each cluster versus subjects in the other clusters. A *t*-test (Matlab software, R2008b) was used to compare the following variables (between paired groups of subjects—DPN versus NoDPN, DPN versus CS, NoDPN versus CS, and between subjects in each cluster versus subjects in the other clusters): velocity (m/s), stride period (s), stride length (m), and stance period (s).

Results

The clinical characteristic of the subjects, the mean and standard deviation of GRF values, joint angles and moments, and time and space parameters according to the clinical subdivision of the subjects are reported in **Tables 1** and **2**.

While considering results of cluster analysis, its classification led to definition of four well-separated clusters when applied to each GRF component: the trunk flexion–extension moment, the knee flexion–extension moment, the ankle abduction–adduction moment, and the ankle internal–external rotation moment.

The estimates of the mean and standard deviation of kinematics, kinetics, and time and space parameters of each family are reported in **Table 3**. The clinical characteristics of each family are reported in **Table 4**.

It should be noticed that 100% of the CS group was completely identified by one cluster that also comprised a smaller percentage of DPN and NoDPN subjects. Meanwhile, subjects of the NoDPN group were equally distributed among three clusters: 40% of NoDPN in the

first cluster (together with CS and DPN), 40% of NoDPN in the second cluster (including only NoDPN), and 40% of NoDPN in the third cluster (together with DPN). Finally, subjects of the DPN group were identified by three clusters; the 11.5% were included in the first cluster (together with CS and NoDPN), the 57.7% in the third cluster (together with NoDPN), and the 30.7% in the fourth cluster, including only DPN. In particular, cluster 1 was characterized by the lowest trunk range of motion on each plane but registered the highest knee and ankle ranges of motion in the sagittal plane. Also, only in this cluster, CS were identified at variance with the other family clusters. Oppositely, cluster 2, which included just NoDPN subjects, showed the highest range of motion at the hip joint both in the sagittal and transverse plane, together with the highest trunk range of motion in the coronal plane. The subjects of this cluster also exhibited the highest mean GRF values.

Cluster 3 registered the highest hip range of motion and mean joint moment together with the lowest trunk range of motion and joint moment in the coronal plane. Furthermore, the subjects in this class registered the highest mean stride and stance period together with the

Table 1.
Clinical Characteristics and Time and Space Parameters (Mean and Standard Deviation) of the Diabetic Neuropathic Group, Diabetic Non-Neuropathic Group, and Control Group[a]

	NoDPN	DPN	CS	DPN versus NoDPN *t*-test (*p*)	DPN versus CS *t*-test (*p*)	NoDPN versus CS *t*-test (*p*)
Subjects (%)	30.3	39.4	30.3	—	—	—
Age (years)	63.8 (5.4)	63.2 (6.0)	59.0 (2.9)	0.70	0.004[b]	0.01[b]
BMI (kg/m^2)	26.3 (2.0)	25.6 (2.9)	24.0 (2.8)	0.4	0.2	0.02[b]
Duration of disease (years)	17.2 (11.7)	22.1(14.3)	—	0.2	—	—
Sex (% of subjects)	F 30 / M 70	F 42.3 / M 57.7	F 35 / M 65	F 0.19 / M 0.07	F 0.31 / M 0.77	F 0.17 / M 0.24
Diabetic retinopathy (% of subjects)	40.0	65.4	—	0.04[b]	—	—
Microalbuminuria (% of subjects)	25.0	34.6	—	0.1	—	—
Peripheral vascular disease (% of subjects)	15.0	19.2	—	0.6	—	—
Autonomic neuropathy (% of subjects)	0	23.0	—	0.02[b]	—	—
Type 1 diabetes (% of subjects)	15.0	42.3	—	0.02[b]	—	—
Type 2 diabetes (% of subjects)	85.0	57.7	—	0.02[b]	—	—
Hemoglobin A1c (%)	7.7 (1.1)	7.6 (1.7)	—	0.1	—	—
Gait velocity (m/s)	1.10 (0.2)	1.10 (0.2)	1.27 (0.1)	0.9	0.003[b]	0.003[b]
Stride period (s)	1.13 (0.1)	1.19 (0.1)	1.1 (0.1)	0.3	0.3	0.4
Stride length (m)	1.23 (0.1)	1.20 (0.2)	1.4 (0.1)	1.0	0.000[b]	0.002[b]
Stance period (s)	0.66 (0.1)	0.73 (0.1)	0.5 (0.1)	0.1	0.000[b]	0.000[b]

[a] The reported *p* values indicate the results of the comparison between the DPN and NoDPN groups, the DPN and CS groups, and the NoDPN and CS groups.
[b] A value of $p < .05$ was considered statistically significant.

Table 2.
Mean and Standard Deviation Results for Ground Reaction Forces, Joint Rotation moments, and Angles Computed over, Respectively, the Stance Phase of Gait and the Gait Cycle on the Diabetic Neuropathic Group, Diabetic Non-Neuropathic Group, and Control Group[a]

	CS	NoDPN	DPN	P1	P2	P3
Forces (%body weight) mean (standard deviation)						
Mediolateral	3.4 (2)	2.9 (1)	1.2 (1)	0.4	0.4	0.4
Vertical	82.4 (25)	89.2 (27)	82.6 (25)	0.4	0.4	0.4
Anterior–Posterior	1.1 (8)	0.3 (6)	0.6 (8)	0.1	0.7	0.7
Moments (%BMI*height) mean (standard deviation)						
Trunk abduction–adduction	5.6 (3)	0.8 (1)	1.1 (1)	0.001[b]	0.001[b]	0.4
Trunk internal–external rotation	0.1 (0.2)	0.9 (0.2)	5.8 (2)	0.4	0.04[b]	0.04[b]
Trunk flexion–extension	-0.9 (0.4)	-0.1(0.3)	-0.1 (0.3)	0.4	0.4	0.7
Hip abduction–adduction	1.2 (3)	2.4 (2)	-4.7(12)	0.4	0.04[b]	0.01[b]
Hip internal–external rotation	3.4 (2)	-14.8 (4)	4.3 (2)	0.0001[b]	0.4	0.0001[b]
Hip flexion–extension	0.7 (0.5)	1.3 (1)	2.0 (0.4)	0.04[b]	0.03[b]	0.4
Knee abduction–adduction	-0.8 (1)	-0.4 (1)	-0.2 (1)	0.3	0.4	0.4
Knee internal–external rotation	0.1 (0.2)	-0.5 (0.2)	3.0 (1)	0.1	0.04[b]	0.01[b]
Knee flexion–extension	1.3 (0.5)	-0.0 (0.2)	2.2 (1)	0.04[b]	0.5	0.001[b]
Ankle abduction–adduction	2.9 (2)	1.8 (2)	0.9 (1)	0.04[b]	0.04	0.02[b]
Ankle internal–external rotation	-0.8 (0.1)	0.9 (1)	3.9 (1)	0.05	0.04	0.001[b]
Ankle flexion–extension	0.4 (0.3)	0.1 (0.1)	-1.2 (0.3)	0.05	0.04	0.04[b]
Rotation (deg) mean (standard deviation)						
Trunk abduction–adduction	0.4 (3)	-1.2 (6)	1.4 (3)	0.001[b]	0.04[b]	0.05
Trunk internal–external rotation	1.3 (4)	-1.9 (3)	0.1 (3)	0.04[b]	0.05	0.05
Trunk flexion–extension	2.5 (1)	-1.4 (0.1)	3.4 (0.6)	0.001[b]	0.002[b]	0.05
Hip abduction–adduction	8.8 (4)	3.7 (4)	-4.7 (9)	0.1	0.04[b]	0.04[b]
Hip internal–external rotation	10.7 (2.9)	-0.0 (0.4)	2.7 (2)	0.001[b]	0.0001[b]	0.04[b]
Hip flexion–extension	15.9 (18)	-0.8 (1)	4.6 (12)	0.00001[b]	0.00001[b]	0.01[b]
Knee flexion–extension	23.1 (20)	9.3 (24)	13.5 (19)	0.05	0.04[b]	
Ankle abduction-adduction	2.4(5)	-6.7(2)	-1.2(6)	0.001[b]	0.05	0.01[b]
Ankle internal-external rotation	1.1(1)	0.0(2)	-4.4(1)	0.05	0.04[b]	0.05
Ankle flexion-extension	2.5(8)	-6.0(7)	-0.5(5)	0.001[b]	0.01[b]	0.0001[b]

[a] Results of one-way ANOVA performed among variables belonging to each group of subjects has been reported. P1, comparison between CS and NoDPN; P2, comparison between CS and DPN; P3, comparison between DPN and NoDPN.
[b] A value of $p < .05$ was considered significant.

Table 3.
Biomechanical Characteristics of Each Cluster: Ground Reaction Forces, Joint Rotation Angles, Joint Rotation moments, Time and Space Parameters, Mean and Standard Deviation of the Diabetic Neuropathic Group, Diabetic Non-Neuropathic Group, and Control Group[a]

	CL 1	CL 2	CL 3	CL 4	CL1–2 (*p*)	CL1–3 (*p*)	CL1–4 (*p*)	CL2–3 (*p*)	CL2–4 (*p*)	CL3–4 (*p*)
Forces (%body weight)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)						
ML	2.90 (1.2)	28.90 (13.1)	1.22 (1.1)	20.42 (1.1)	0.04[b]	0.04[b]	4E-16[b]	4E-03[b]	0.9	4E-04[b]
V	89.20 (27.2)	161.30 (17.1)	82.6(24.8)	149.7 (24.9)	2E-10[b]	0.8	1E-10[b]	1E-03[b]	0.08	1E-04[b]
AP	0.33 (6.5)	7.79 (59.2)	0.57 (7.6)	3.55 (7.5)	1E-10[b]	0.4	0.5	0.001[b]	0.001[b]	0.7
Moments (%BMI*H)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)						
Trunk AA	5.46 (2.3)	-2.05 (0.7)	1.73 (2.5)	-2.23 (0.6)	0.7	0.3	0.7	0.9	1E-03[b]	0.001[b]
Trunk IE	-0.08 (0.4)	-0.14 (0.3)	-0.83 (0.6)	-0.11 (0.1)	0.8	0.6	0.8	0.6	9E-03[b]	0.05
Trunk FE	-0.04 (0.2)	-0.06 (0.2)	-0.39 (0.3)	0.05 (0.1)	0.6	0.9	0.6	2E-03[b]	1E-03[b]	8E-05[b]
Hip AA	5.91 (2.3)	4.15 (1.8)	26.99 (8.9)	0.65 (0.5)	0.2	4E-08[b]	9E-03[b]	0.001[b]	4E-08[b]	1E-10[b]
Hip IE	-2.66 (2.2)	-1.67 (4.1)	-5.15 (14.5)	-0.69 (1.8)	0.8	0.7	0.4	0.05	0.2	0.1
Hip FE	-5.42 (2.2)	-5.54 (1.5)	-21.99 (6.5)	-2.71 (0.9)	0.7	0.9	0.03	0.4	2E-04[b]	0.04[b]
Knee AA	-0.43 (0.98)	-0.47 (1.7)	-0.98 (1.3)	0.14 (0.6)	0.7	0.1	0.7	0.1	0.04[b]	0.04[b]
Knee IE	-0.76 (0.5)	-1.27 (0.4)	-0.79 (0.5)	0.07 (0.1)	0.7	0.7	0.1	0.04[b]	0.1	0.04[b]
Knee FE	-0.12 (0.3)	0.28 (0.5)	-1.03 (0.8)	0.01 (0.1)	0.02[b]	0.09	7E-04[b]	0.08	0.001[b]	0.007[b]
Ankle AA	2.91 (2.3)	2.56 (2.3)	0.67 (0.6)	0.96 (1.0)	0.7	0.5	0.3	0.0002[b]	0.7	0.0002[b]
Ankle IE	0.59 (0.6)	0.76 (0.8)	-0.34 (0.6)	0.69 (0.5)	0.4	0.3	0.04[b]	0.3	0.004[b]	0.7
Ankle FE	-0.17 (0.1)	-0.13 (0.1)	-0.15 (0.2)	0.01 (0.1)	0.1	0.2	0.5	0.3	0.0001[b]	0.0001[b]
Rotations (deg)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)						
Trunk AA	0.71 (1.6)	-0.60 (2.9)	-1.22 (5.0)	3.21 (2.4)	0.4	0.3	0.07	0.1	0.2	0.04[b]
Trunk IE	-0.36 (2.3)	1.29 (2.6)	-0.21 (4.4)	-0.58 (1.5)	0.4	0.9	0.7	0.8	0.04[b]	0.5
Trunk FE	-0.49 (3.0)	-1.29 (3.5)	2.31 (5.8)	1.62 (0.5)	0.5	0.5	0.8	0.3	0.6	0.6
Hip AA	2.98 (3.1)	1.14 (4.4)	1.33 (5.0)	-1.81 (2.2)	0.1	0.1	0.5	0.1	0.3	0.5
Hip IE	-0.16 (4.5)	-0.42 (4.9)	-2.60 (5.2)	-1.00 (2.7)	0.4	0.09	0.9	0.2	0.1	0.9
Hip FE	12.16 (3.9)	34.76 (4.9)	-1.13 (12.2)	13.5 (2.5)	0.02[b]	0.8	0.01[b]	0.08	1E-03[b]	0.9
Knee FE	21.57 (6.6)	62.92 (6.5)	13.84 (12.0)	23.22 (13.9)	0.04[b]	0.8	0.04[b]	0.02[b]	1E-03[b]	0.3
Ankle AA	-1.12 (4.4)	-3.05 (2.6)	-1.20 (7.4)	-18.89 (20.8)	0.4	0.7	0.02[b]	0.5	0.9	0.6
Ankle IE	-1.26 (1.9)	-0.99 (1.7)	-6.96 (10.1)	28.21 (18.5)	0.03[b]	0.6	0.02[b]	0.8	0.006[b]	0.8
Ankle FE	-1.78 (4.1)	7.49 (13.9)	0.28 (24.7)	39.81 (68.6)	0.6	0.04[b]	0.7	0.6	0.01[b]	0.01[b]

[a] Results of one-way ANOVA performed among variables belonging to each cluster (CL) have been reported. SD, standard deviation; ML, mediolateral; V, vertical; AP, anterior–posterior; AA, abduction–adduction; IE, internal–external rotation; FE, flexion–extension.
[b] A value of $p < .05$ was considered significant.

Table 4.
Clinical Characteristics and Time and Space Parameters (Mean and Standard Deviation) within Each Cluster[a]

	CL1	CL2	CL3	CL4	CL1–2 (*p*)	CL1–3 (*p*)	CL1–4 (*p*)	CL2–3 (*p*)	CL2–4 (*p*)	CL3–4 (*p*)
Subjects (% of subjects within the total number of subjects)	47	12	35	12	0	0.08	0	0.001[b]	0.5	0.001[b]
CS (% of CS within the specific cluster)	64	0	0	0	0	6E-04[b]	6E-04[b]	—	—	—
NoDPN (% of NoDPN within the specific cluster)	26	100	35	0	1E-04[b]	0.2	0.05	7E-04[b]	0	0.03[b]
Age (years)	62.9 (5)	61.8 (4)	65.1 (6)	65.0 (6)	0.4	0.003[b]	0.003[b]	0.003[b]	0.003[b]	0.4
BMI (kg/m^2)	25.7 (2)	24.5 (2)	25.8 (3)	25 (4)	0.4	0.4	0.4	0.4	0.4	0.4
Duration of disease (years)	19.8 (14)	12.0 (11)	26.0 (10)	15.5 (9)	0.08	0.8	0.1	0.009[b]	0.1	0.01[b]
Sex (F/M, % of subjects within the specific cluster)	39 61	50 50	26 74	50 50	0.3 0.3	0.2 0.2	0.3 0.3	0.5 0.5	0.1 0.1	0.1 0.1
DPN (% of DPN subjects within the specific cluster)	10	0	65	100	0.2	0	0	7E-04[b]	0	0.03[b]
Diabetic retinopathy (% of subjects within the specific cluster)	13	25	52	50	0.8	9E-04	0.01[b]	0.09	0.1	0.4
Microalbuminuria (% of subjects within the specific cluster)	10	0	22	25	0.2	0.1	0.1	0.07	0.06	0.4
Peripheral vascular disease (% of subjects within the specific cluster)	3	25	4	50	0.02[b]	0.4	2E-04[b]	0.04[b]	0.1	0.001[b]
Autonomic neuropathy (% of subjects within the specific cluster)	0	0	22	50	—	0.003[b]	0	0.07	0.01[b]	0.06
Type 1 (% of subjects within the specific cluster)	13	25	30	37.5	0.8	0.06	0.05	0.4	0.3	0.4
Type 2 (% of subjects within the specific cluster)	23	75	70	62.5	0.002[b]	3E-04[b]	0.01[b]	0.4	0.3	0.3
Hemoglobin A1c (%)	7.9 (1)	7.8 (1)	7.5 (1)	7.0 (1)	0.3	0.9	0.3	0.3	0.9	0.2
Mean velocity (m/s)	1.13 (0.2)	1.04 (0.1)	0.95 (0.3)	1.10 (0.05)	0.5	0.6	0.5	0.1	0.4	0.8
Stride period (s)	1.14 (0.1)	1.10 (0.1)	1.18 (0.1)	1.10 (0.04)	0.8	0.9	0.6	0.4	0.03[b]	0.8
Stride length (m)	1.30 (0.1)	1.10 (0.1)	1.15 (0.2)	1.40 (0.3)	0.8	0.1	0.7	0.2	0.01[b]	0.8
Stance period (s)	0.60 (0.1)	0.60 (0.03)	0.75 (0.1)	0.60 (0.04)	0.8	0.1	0.1	0.4	0.7	0.9

[a] The percentage of subjects displaying each clinical characteristic as given in the table represents the percentage of subjects within the considered cluster (CL). Results of the comparison among the clusters' clinical characteristics and time and space parameters (*t*-test, *z*-test, R Statistic software) have been reported in terms of *p* value.
[b] A value of $p < .05$ was considered statistically significant.

lowest gait velocity (**Table 4**). In particular, it included both DPN and NoDPN subjects. Finally, cluster 4, which included only DPN subjects, was identified. This was characterized by excessive GRF values as well as the highest ankle range of motion in each plane. The longest stride length was also registered in this cluster, together with the lowest hip and knee joint moments in each plane.

Discussion

In this article, we have demonstrated, for the first time, the advantage of describing DPN and NoDPN biomechanical alteration starting from their gait pattern characteristics by means of cluster analysis. Rather than grouping subjects according to their metabolic control, duration of the disease, gender, BMI, and age, we chose to stratify subjects according to their biomechanical characteristics. *K*-means cluster analysis was chosen using gait parameters only, and it produced four well-separated clusters. This was successfully obtained by seven variables. The effectiveness of such a classification was confirmed by the one-way ANOVA results, which revealed statistically significant differences among variables of different clusters. Furthermore, not only the variables that defined well-separated clusters but also other variables displayed important differences among subjects of different clusters (see **Table 3**). Four families were identified, thus accomplishing the primary purpose of this study, which was to verify the presence of more than one gait pattern in DPN and NoDPN subjects. The need for such data was evident given the limitation of procedures used for data analysis undertaken in previous research. As such, until now, it was difficult to appreciate differences within DPN or NoDPN gait patterns occurring during the stride cycle, thus allowing the possibility of statistical errors and undermining the validity of the comparison. It has been stated that, if different movement strategies exist, then statistical power will be lowered, and there is a greater likelihood of falsely supporting the null hypothesis in group comparison statistical procedures.[31,32] In this article, we chose a technique able to detect the presence of other subgroups in data, thus allowing the possibility of making valid comparisons. Hence our results showed the presence of seven variables that defined three different patterns of gait either in DPN or NoDPN subjects. Moreover, most of the DPN subjects were included in cluster 3, which contained both DPN and NoDPN subjects. This is an important result if we consider that the majority of the studies analyzed averaged data obtained by DPN subjects separated from NoDPN subjects, hypothesizing a single gait pattern for each pathologic population. It should be further mentioned that, in cluster 3, an important presence of diabetes complications was noticed together with a long duration of the disease. Gait parameters of its subjects showed an important role played by the hip joint with respect to the ankle joint. This seems to suggest a changing in the walking strategy from the ankle to the hip joint[13,33] in relation to a long duration of the disease and the presence of complications. This was also confirmed by previous research that assessed a change in the walking strategy as one of the consequences of diabetes.[13–33] Nevertheless, a nice correlation has already been reported between diabetes complications and the presence of alteration in the gait pattern of a group of DPN and NoDPN subjects.[15]

In analyzing the results of the present classification, the identification of cluster 1 should be mentioned for being characterized either by pathologic subjects or CS. This shows how a sample of pathologic subjects can display a gait pattern similar to CS. With respect to the other families, this cluster displayed a lower number of subjects with diabetes complications, thus confirming their important role in altering the gait pattern of diabetes patients. This was also confirmed by the higher GRFs registered in clusters 2 and 4 that included a larger proportion of vasculopathic subjects. Indeed, it has been demonstrated that motor dysfunction occurs in peripheral vascular disease subjects, where inadequate blood flow to the lower limb may result in reduction of physiological capacity, lower limb mobility, walking performance, physical activity levels, and decreased health-related quality of life.[15,34] It should be further mentioned that, although the proposed methodology allowed identifications of more than one gait pattern in the studied subjects, two families (clusters 2 and 4) including, respectively, only DPN and NoDPN subjects were identified. This demonstrated the possibility of distinguishing gait patterns of DPN and NoDPN subjects from those of CS exclusively by means of gait data. These families were characterized by excessive GRF components that confirmed the important role played by either glycosylation and diabetes complications. Glycosylation alters the gait pattern in diabetic subjects by reducing the ability to accomplish shock absorption during gait. This was also reported in previous literature, even though classical statistic methodologies were adopted in assessing diabetic and neuropathic gait alterations.[1,8–15,33–36]

Despite its potential usefulness, cluster analysis has not been widely used in clinical biomechanics. This may be due to unfamiliarity and/or concerns among biomechanists

regarding procedural problems associated with this technique.[31] One of the main drawbacks of the cluster analysis method is that it is sensitive to variations in the gait trials, especially where reproducibility may be difficult to attain,[16,18] particularly in the neuropathic subjects. The trials of a new subject may not be grouped together but, rather, may be located anywhere in the different clusters, and at the same time, adding new subjects can perturb the clustering. It is also possible that the trials of a diabetes subject may be located in one of the CS clusters. Inversely, the trials of a control individual could end up within the cluster of the diabetes subjects. This study was susceptible to this type of perturbation. However, considering that the principal aim of the present project was observing the natural grouping of subjects in order to display the presence of more than one gait pattern in each of the observed populations, this type of perturbation does not represent a critical source of error. Furthermore, the variability expressed by the standard deviation in the kinematics and kinetics parameters was always larger in the DPN group than in the CS group, thus emphasizing the possibility of different families of neuropathic gaits, but not CS one. However, this possibility was in agreement with the aims of this article.

In conclusion, this methodology clearly provided results that could be hidden if only a group ensemble is presented as DPN or NoDPN for any biomechanic variable.

Conclusions

Biomechanical issues involved in the various phases of treating foot problems in people with diabetes[1–15] have been recognized to be clinically relevant.

This article proposed a method that allows automatic detection of diabetic gait alterations and enabled the identification of a group of NoDPN subjects that displayed similar alterations to DPN subjects. Such a result could have important clinical outcomes. Indeed, this family of subjects was characterized by long disease duration and several diabetes complications. This seems to suggest that the presence of these characteristics has an important role in altering the gait pattern of diabetic subjects. Furthermore, the family including only NoDPN subjects displayed higher GRF components, thus highlighting how diabetes reduces the ability to accomplish shock absorption during gait.

References:

1. Van Schie CH. A review of the biomechanics of the diabetic foot. Int J Low Extrem Wounds. 2005;4(3):160–70.
2. Greene DA, Sima AA, Stevens MJ, Feldman EL, Lattimer SA. Complications: neuropathy, pathogenetic considerations. Diabetes Care. 1992;15(12):1902–25.
3. Boulton AJ, Gries FA, Jervell JA. Guidelines for the diagnosis and outpatient management of diabetic peripheral neuropathy. Diabet Med. 1998;15(6):508–14
4. Boucher P, Teasdale N, Courtemanche R, Bard C, Fleury M. Postural stability in diabetic polyneuropathy. Diabetes Care. 1995;18(5):638–45.
5. Geurts AC, Mulder TW, Nienhuis B, Mars P, Rijken RA. Postural organization in patients with hereditary motor and sensory neuropathy. Arch Phys Med Rehabil. 1992;73(6):569–72.
6. Simoneau GG, Ulbrecht JS, Derr JA, Becker MB, Cavanagh PR. Postural instability in patients with diabetic sensory neuropathy. Diabetes Care. 1994;17(12):1411–21.
7. Bloem BR, Allum JH, Carpenter MG, Honegger F. Is lower leg proprioception essential for triggering human automatic postural responses? Exp Brain Res. 2000;130(3):375–91.
8. Katoulis EC, Ebdon-Parry M, Hollis S, Harrison AJ, Vileikite L, Kulkarni J, Boulton AJ. Postural instability in diabetic neuropathic patients at risk of foot ulceration. Diabet Med. 1997;14(4):296–300.
9. Dingwell JB, Ulbrecht JS, Boch J, Becker MB, O'Gorman JT, Cavanagh PR. Neuropathic gait shows only trends towards increased variability of sagittal plane kinematics during treadmill locomotion. Gait Posture.1999;10(1):21–9.
10. Dingwell JB, Cusumano JP, Sternad D, Cavanagh PR. Slower speeds in patients with diabetic neuropathy lead to improved local dynamic stability of continuous overground walking. J Biomech. 2000;33(10):1269–77.
11. Dingwell JB, Cavanagh PR. Increased variability of continuous overground walking in neuropathic patients is only indirectly related to sensory loss. Gait Posture. 2001;14(1):1–10.
12. Courtemanche R, Teasdale N, Boucher P, Fleury M, Lajoie Y, Bard C. Gait problems in diabetic neuropathic patients. Arch Phys Med Rehabil. 1996;77(9):849–55.
13. Mueller MJ, Sinacore DR, Hoogstrate S, Daly L. Hip and ankle walking strategies, effect on Peack plantar pressures and implications for neuropathic ulceration. Arch Phys Med Rehabil. 1994;75(11):1196–200.
14. Sawacha Z, Cristoferi G, Guarneri G, Corazza S, Donà G, Denti P, Facchinetti A, Avogaro A, Cobelli C. Characterizing multisegment foot kinematics during gait in diabetic foot patients. J Neuroeng Rehabil. 2009;6:37.
15. Sawacha Z, Gabriella G, Cristoferi G, Guiotto A, Avogaro A, Cobelli C. Diabetic gait and posture abnormalities: a biomechanical investigation through three dimensional gait analysis. Clin Biomech. 2009;24(9):722–8.
16. Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate data analysis with readings. 2nd ed. New York: Macmillan; 1990.
17. Vardaxis VG, Allard P, Lachance R, Duhaime M. Classification of able-bodied gait using 3-D muscle powers. Hum Mov Sci. 1998:17(1):121–36.
18. Watelain E, Barbier F, Allard P, Thevenon A, Angué JC. Gait pattern classification of healthy elderly men based on biomechanical data. Arch Phys Med Rehabil. 2000;81(5):579–86.
19. White SG, McNair PJ. Abdominal and erector spinae muscle activity during gait: the use of cluster analysis to identify patterns of activity. Clin Biomech (Bristol, Avon). 2002;17(3):177–84.

20. Vannozzi G, Cereatti A, Mazzà C, Benvenuti F, Della Croce U. Extraction of information on elder motor ability from clinical and biomechanical data through data mining. Comput Methods Programs Biomed. 2007;88(1):85–94.

21. Tan PN, Steinbach M, Kumar V. Introduction to data mining. Massachusetts: Addison-Wesley; 2006.

22. Altman DG. Practical statistics for medical research. London: Chapman & Hall; 1991.

23. American Diabetes Association, American Academy of Neurology. Consensus statement, report and recommendations of the San Antonio conference on diabetic neuropathy. Diabetes Care. 1988;11(7):592–7.

24. Bourdiol JR. Pied et statique. Paris: Maisonneuve;1980.

25. Ashford RL, Shippen J. Leg length measurement: clinical versus mathematical modelling. Foot. 2003;13(3):174–8.

26. Cappozzo A, Catani F, Croce UD, Leardini A. Position and orientation in space of bones during movement: anatomical frame definition and determination. Clin Biomech (Bristol, Avon). 1995;10(4):171–8.

27. Leardini A, Sawacha Z, Paolini G, Ingrosso S, Nativo R, Benedetti MG. A new anatomically based protocol for gait analysis in children. Gait Posture. 2007;26(4):560–71.

28. Wu G, Siegler S, Allard P, Kirtley C, Leardini A, Rosenbaum D, Whittle M, D'Lima DD, Cristofolini L, Witte H, Schmid O, Stokes I, Standardization and Terminology Committee of the International Society of Biomechanics. ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion--part I: ankle, hip and spine. J Biomech. 2002;35(4):543–8.

29. Della Croce U, Leardini A, Chiari L, Cappozzo A. Human movement analysis using stereophotogrammetry. Part 4: assessment of anatomical landmark misplacement and its effects on joint kinematics. Gait Posture. 2005;21(2):226–37.

30. Everitt BS, Landau S, Leese M. Cluster analysis. 4th ed. New York: Oxford University Press; 2001.

31. Ball KA, Best RJ. Different centre of pressure patterns within the golf stroke I: Cluster analysis. J Sports Sci. 2007;25(7):757–70.

32. Bates BT. Single-subject methodology: an alternative approach. Med Sci Sports Exerc. 1996;28(5):631–8.

33. Giacomozzi C, Caselli A, Macellari V, Giurato L, Lardieri L, Uccioli L. Walking strategy in diabetic patients with peripheral neuropathy. Diabetes Care. 2002;25(8):1451–7.

34. Crowther RG, Spinks WL, Leicht AS, Quigley F, Golledge J. Lower limb movement variability in patients with peripheral arterial disease. Clin Biomech (Bristol, Avon). 2008;23(8):1080–5.

35. D'Ambrogi E, Giurato L, D'Agostino MA, Giacomozzi C, Macellari V, Caselli A, Uccioli L. Contribution of plantar fascia to the increased forefoot pressures in diabetic patients. Diabetes Care. 2003;26(5):1525–9.

36. Akashi PM, Sacco IC, Watari R, Hennig E. The effect of diabetic neuropathy and previous foot ulceration in EMG and ground reaction forces during gait. Clin Biomech (Bristol, Avon). 2008;23(5):584–92.