

An index for quantifying deviations from normal gait

L.M. Schutte^{a,b}, U. Narayanan^{a,b}, J.L. Stout^a, P. Selber^c, J.R. Gage^{a,b},
M.H. Schwartz^{a,b,*}

^a Gillette Children's Specialty Healthcare, St. Paul, MN, USA

^b Department of Orthopaedic Surgery, University of Minnesota, Minneapolis, MN, USA

^c AACD, Sao Paulo, SP, Brazil

Received 28 January 1999; accepted 7 October 1999

Abstract

A method is derived to calculate the amount by which a subject's gait deviates from an average normal profile, and to represent this deviation as a single number. The method uses principal component analysis to derive a set of 16 independent variables from 16 selected gait variables. The sum of the square of these 16 independent variables is interpreted as the deviation of the subject's gait from normal. Statistical tests of the method's validity and an initial demonstration of its clinical utility are included. It is found that using this index, increasing clinical involvement corresponds to increasing index score. © 2000 Published by Elsevier Science B.V. All rights reserved.

Keywords: Gait analysis; Multivariate analysis; Cerebral palsy; Human locomotion

1. Introduction

Gait analysis provides an effective tool for evaluating and quantifying the effects of a surgical intervention or other treatment on a patient's gait [1]. An experienced clinician can make a subjective evaluation as to whether a patient's gait has become more normal following intervention and, with the gait analysis data, quantify the specific features of the gait that have changed. However, even with the assistance of gait analysis, objectively quantifying the degree to which a patient's gait has improved following an intervention remains difficult.

Most research studies using gait analysis have relied on comparisons of a limited number of specific gait characteristics to evaluate the effects of surgical procedures [2–6]. Such studies ignore the high degree of correlation that exists between various aspects of an individual's gait. Peak knee flexion during gait, for example, has been previously shown to be highly dependent on walking speed [7]. Similarly, hip flexion depends on the position of the pelvis, knee flexion and ankle dorsiflexion both depend on the orientation of

the shank, and hip rotation and foot progression angle are highly coupled. In only a few cases have techniques, based on multivariate statistics that attempt to account for such correlation, been used to describe gait patterns. For example, Mah et al. [8] apply principal component analysis to time series functions of eight measured kinematic variables in order to determine three time series principle components describing over 90% of the information in the gait data. In addition, Loslever et al. [9–11] use a correspondence factor analysis method to describe different walking patterns within a group of normal subjects.

To accurately evaluate the extent of gait deviations from normal gait, or to assess the changes in a gait resulting from a specific treatment, it is important to consider not only how each feature of the gait pattern has changed but also how the relationship between the features changed. To evaluate whether a specific gait variable is normal, abnormal, or improved following treatment, the natural correlation that exists between gait variables must be determined. For this reason multivariate statistical techniques are used to develop a measure of how closely an individual gait pattern approaches normal. This 'closeness' is referred to as the normalcy index.

* Corresponding author.

2. Methodology

In simplest terms, the normalcy index can be considered a measure of the distance between the set of discrete variables describing a patient's gait pattern and the average of those variables in persons with no gait abnormalities. Because the discrete gait variables are correlated, a simple distance calculation between the original discrete variables and the mean value that is expected for a normal gait pattern would not be expected to accurately represent the degree of gait abnormality (Fig. 1). However, using techniques from multivariate statistics it is possible to uncorrelate the discrete variables and calculate the distance in a new uncorrelated coordinate system. The normalcy index is the square of this uncorrelated distance. The procedure used to calculate the index is outlined below and illustrated in Fig. 2.

2.1. To calculate the normalcy index

Choose N discrete gait variables. Let these variables be represented by x_j , $j = 1, N$. The choice of variables is inherently arbitrary. For this project a group of experienced clinicians were polled to arrive at an appropriate set of variables. Variables were selected that the clinicians felt correlated closely with particular gait problems. For example, clearance of the foot in swing is an

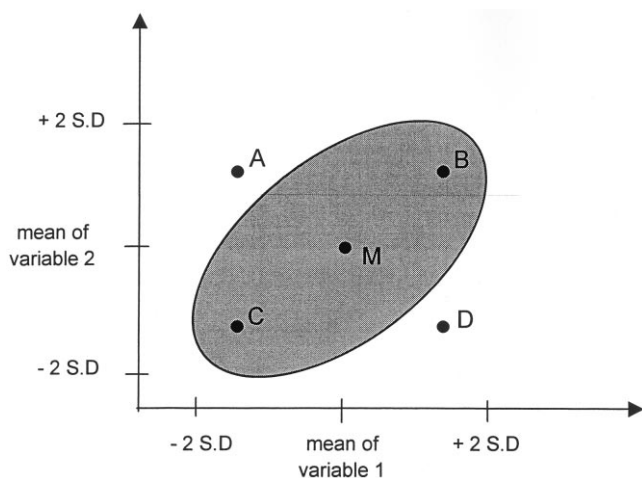


Fig. 1. Illustration of how correlation influences the interpretation of normalcy for two correlated variables. The ellipse represents the area within two S.D.s of the mean of two hypothetical normally distributed variables that have a strong correlation. All combinations of variables within the ellipse are within two S.D.s of the mean combination M . Each of the points represented by A , B , C , and D is a pair of data points within two S.D.s of the individual means of each of the two variables. If each variable were considered individually all four would be considered equally likely. However, A and D are outside of the two S.D. region while B and C are within. That is, because of the correlation between the two variables A and D represent a combination of the two variables that are less likely to occur than the combinations represented by B and C .

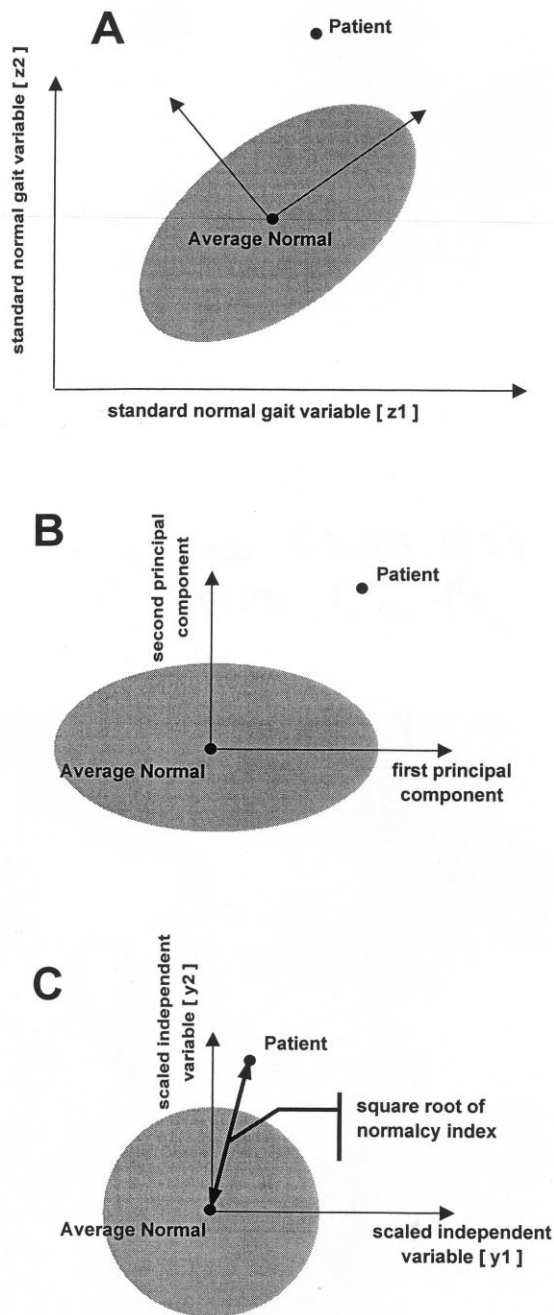


Fig. 2. A hypothetical two variable demonstration of the method used to calculate the normalcy index. (A) Combinations of the two correlated gait variables that are within two S.D.s of the mean normal combination are contained within an ellipse situated at an angle with respect to the coordinate axes. A specific patient is represented by a particular combination of these two variables. (B) The principal components (eigenvectors) are coordinate axes aligned with the major and minor axes of the two S.D. ellipse. The combination of principal components (principal component scores) for the patient can be seen in the rotated coordinate system. (C) The scaled uncorrelated variables y_1 and y_2 are shown. The ellipse, plotted in these uncorrelated and scaled coordinates, becomes a circle. The distance from the transformed point representing the specific patient to the origin provides a meaningful measure of how far the patient's combination of standard normal gait variables is from the average combination of these variables. The normalcy index is the square of this distance.

important pre-requisite of normal gait and is closely correlated with the magnitude and timing of peak knee flexion as well as the magnitude of dorsiflexion during swing.

1. Let μ_j , σ_j be the mean and S.D. of x_j measured on M normal subjects.
2. Let z_j be a set of standardized discrete gait variables (zero mean, unit S.D.) defined by,

$$z_j = (x_j - \mu_j)/\sigma_j, \quad j = 1, N \quad (1)$$

3. Note that this standardization is required so that the different units of the discrete variables do not skew the results. These variables are correlated with one another (Fig. 2A).
4. Calculate the covariance matrix, C_{ij} , for the N standardized discrete variables. The covariance matrix is an $N \times N$ matrix with the ij position occupied by the correlation coefficient between the i th and the j th variable.
5. Calculate the N eigenvalue–eigenvector pairs, λ_i – \mathbf{e}_i , of the covariance matrix. The eigenvectors can be expressed as:

$$\mathbf{e}_i = \sum_{j=1}^N \alpha_j^i x_j \quad (2)$$

where x_j are the original gait variables and α_j^i are relative weights determined solely by the orthonormality constraint. The eigenvectors form a set of N independent/uncorrelated variables, called principal components, that are linear combinations of the original N discrete variables [12]. The eigenvectors may be thought of as orthogonal coordinate axes spanning an N dimensional space such that

$$\mathbf{e}_i \cdot \mathbf{e}_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \quad (3)$$

It is worth noting that an infinite number of such coordinate systems can be derived via a rotation of the original coordinate system (Fig. 2B). This fact imposes some logical imperatives on the distance measure proposed for index (step 10). The average normal subject lies at the origin of this space.

6. Define a new set of scaled independent/uncorrelated variables, y_i , such that each y_i has equal variance over the population of M normal subjects (Fig. 2C). This is accomplished through division by the square root of the corresponding eigenvalues.

$$y_i = \left(\frac{1}{\lambda_i}\right) \sum_{j=1}^N \alpha_j^i z_j \quad (4)$$

By applying this scaling the magnitude of variation inherent in certain variables is accounted for. In

other words, if one of the original N variables (or some linear combination of the N variables) has a large variation within the normal population, then a large deviation from the average value of that variable will not count excessively against the ‘normalcy’ of a subject with some pathology. Arguments against such a scaling are not without merit. It could be argued, for example, that the eigenvectors associated with small eigenvalues represent variable combinations that may be small random fluctuations and should not be magnified through division by their eigenvalue.

7. Let a subject’s gait be represented by the same N discrete variables \tilde{x}_j , $j = 1, N$.
8. Scale \tilde{x}_j by the average normal gait variables as follows:

$$\tilde{z}_j = (\tilde{x}_j - \mu_j)/\sigma_j, \quad j = 1, N, \quad (5)$$

where μ_j and σ_j are the mean and S.D. from the normal population (step 2 above).

9. Apply the transformation of Eq. (4) to the patient’s gait variables:

$$\tilde{y}_i = \left(\frac{1}{\lambda_i}\right) \sum_{j=1}^N \alpha_j^i \tilde{z}_j \quad (6)$$

where again λ_i and α_j^i are derived from the normal population (step 5 above).

10. Find the square of the Euclidean length of \tilde{y}_i for a given subject

$$d = \sum_{i=1}^N \tilde{y}_i^2 \quad (7)$$

This number, d , represents the square of the distance of an individual patient’s data from the normal mean in the new uncorrelated coordinate system. This number is referred to hereafter as the ‘normalcy index’.

The gait of 71 patients and 24 controls with no known gait abnormalities seen in the Gillette Children’s Motion Lab between 1993 and 1996 was measured. The age range of the patients pre-operatively was 7 years 2 months to 19 years 10 months (average 10 years 4 months), while the normals ranged in age from 4 years 11 months to 17 years 7 months (average 10 years 6 months). The patients all had a diagnosis of cerebral palsy although their level of involvement ranged from mild hemiplegia to severe quadriplegia. The hemiplegic patients were classified into Type I, Type II, Type III and Type IV based on the classification scheme described in Winters et al. [6].

Three-dimensional positions of 13 reflective markers attached to the subjects were tracked using an optical motion measurement system (VICON, Oxford Metrics, Oxford, UK) as each subject walked at a self-selected speed. The VICON Clinical Manager (VCM) software was used to calculate three-dimensional joint angles

from each subject's measurements and these joint angles were reported at intervals of 2% of the total gait cycle. VCM calculates joint angles as ordered rotations between anatomically aligned reference frames associated with adjacent body segments [1,13].

Several walking trials were collected for each subject. From these, one representative gait cycle on the right and one representative gait cycle on the left were selected. Sixteen discrete variables were determined from the kinematic measurements for each gait cycle. These discrete variables are: time to foot off as a percent of the total gait cycle time, walking speed normalized by leg length, cadence, mean pelvic tilt, pelvic range of motion in the sagittal plane, mean pelvic rotation, minimum hip flexion, total range of hip flexion–extension, peak abduction in swing, mean hip rotation in stance, knee flexion at initial contact, time to peak knee flexion, total range of knee flexion–extension, peak dorsiflexion in stance, peak dorsiflexion in swing and mean foot progression angle in stance.

The normalcy index, based on these 16 discrete variables, was calculated for each gait cycle of both individuals with cerebral palsy and the control subjects with no gait abnormalities. Separate measures for the right and left gait cycle for each subject were obtained.

3. Results

A number of analyses were conducted to gain confidence in the index and establish its clinical utility. The objective in selecting the 16 discrete variables used to calculate the index was to have a finite set of variables that well described an individual's gait pattern. Variables generally considered to be clinically important

were chosen. Still the process of selecting the set of variables to use was subjective. To evaluate the consequences of the choices the index was recalculated leaving out one variable each time. Using the measurements from all 95 subjects (subjects with gait pathologies and controls), the correlation between the original index and indexes calculated with one variable omitted were determined. The normalcy index appears to be relatively insensitive to its exact composition. That is, none of the 16 gait variables chosen for inclusion in the index seem to dominate the final value (Table 1). In all cases the correlation coefficient between the index calculated with all 16 variables and the index calculated with 15 of the 16 variables was found to be greater than 0.95. In 13 of the 16 cases the correlation coefficient calculated was greater than 0.99. The lowest correlation (0.96) was between the indexes calculated with and without the time to peak knee flexion. This indicates that there may be a smaller set of variables that could also adequately describe the gait. For other analyses that are computationally intensive, such as forward dynamic simulations or dynamic stability analyses, finding this minimal variable set may be of great value. For the purposes of this study it was deemed to be of less importance since the computational burden of 16 variables was minimal.

Next, the distribution of the control subjects' indexes was compared to the χ^2 distribution (Fig. 3). The distribution of a sum of squares of N normally distributed quantities is represented by a χ^2 distribution with $N-1$ degrees of freedom [14,15]. The control subjects' normalcy indexes were thus expected to follow a χ^2 distribution of 15 degrees of freedom. Such a χ^2 distribution has a maximum peak at 13 and 99.5% of the values are below 33. The control subjects' indexes all were found to be between 5 and 30 with the greatest number (16) between 10 and 15. The distribution of the control subjects' indexes was also seen to be skewed slightly towards the right (i.e. the distribution to the right of the peak is more spread out than to the left) similar to the χ^2 distribution.

It is desirable that the index be able to distinguish between patients with different severity of gait abnormality. To test this, the average and range of index values were determined for each of the diagnosis groups: controls, type I hemiplegia, type II hemiplegia, type III hemiplegia, type IV hemiplegia, diplegia and quadriplegia. The average and ranges for the hemiplegia patients were calculated with the affected and unaffected side considered separately and averaged. It is expected that, on average, the more severe diagnoses would correspond to more gait abnormalities and thus result in higher normalcy indexes. That is the index values, on average, should be higher for patients with a diagnosis of type II hemiplegia than for patients with a diagnosis of type I hemiplegia; higher for patients with Type III hemiplegia than for patients with Type II

Table 1

Excluded parameter	Correlation ^a
Time of toe off	0.991
Walking speed	0.998
Cadence	0.999
Mean pelvic tilt	0.996
Range of pelvic tilt	0.987
Mean pelvic rotation	0.995
Minimum hip flexion	0.997
Range of hip flexion	0.998
Peak abduction in swing	0.998
Mean hip rotation in stance	0.996
Knee flexion at initial contact	0.986
Time of peak knee flexion	0.960
Range of knee flexion	0.998
Peak dorsiflexion in stance	0.997
Peak dorsiflexion in swing	0.994
Mean foot progression angle	0.998

^a Correlation between index calculated with all 16 parameters and index calculated with one excluded parameter.

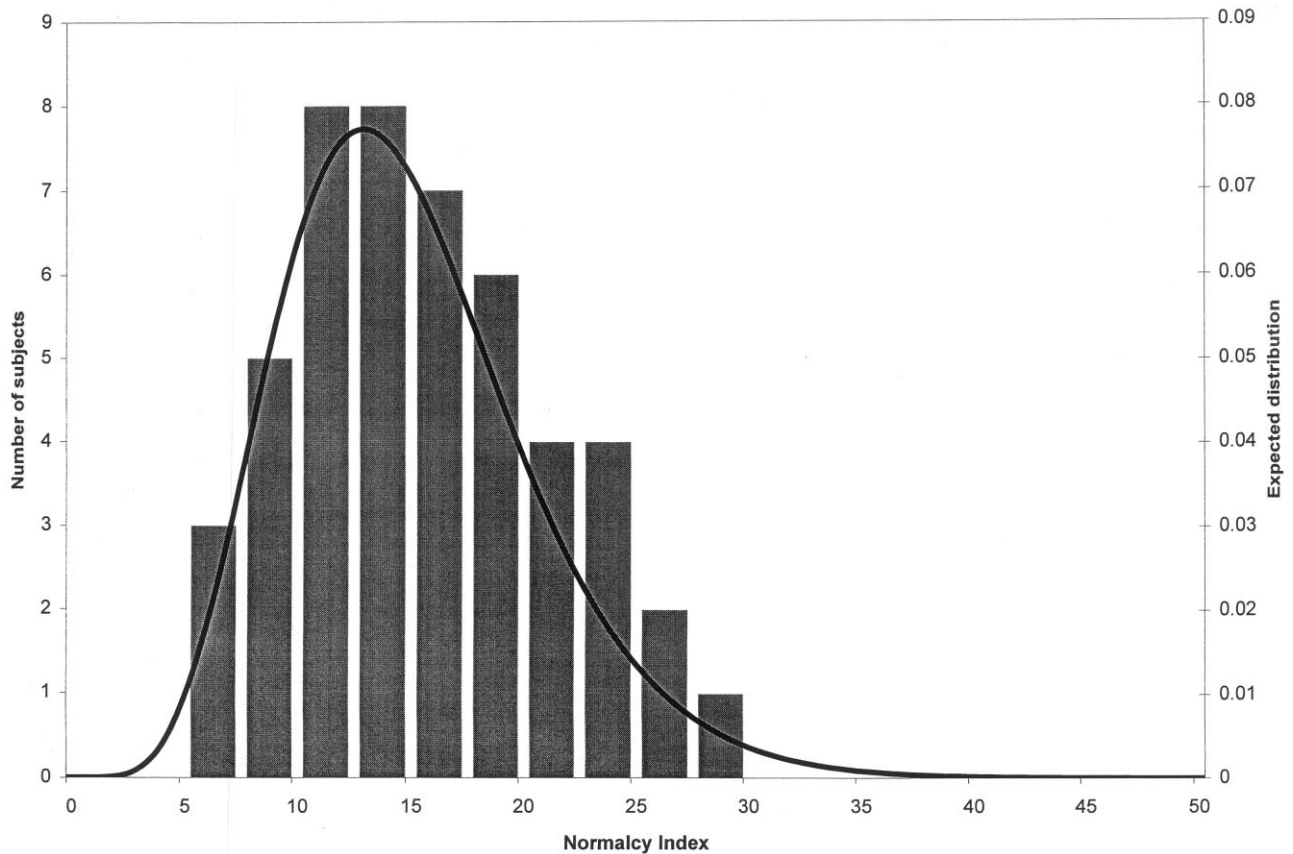


Fig. 3. Actual and expected distributions of normalcy index values for subjects without identified gait abnormalities. The distribution of the indexes calculated individually for the right and left side of the 24 control subjects is represented as a histogram. The expected distribution, a χ^2 distribution with 15 degrees of freedom (solid line), is shown for comparison. The actual and expected distributions compare favorably.

Table 2

Subjects	N	Mean normalcy index (range) ^a	Mean unaffected side only (range)	Mean affected side only (range)
Normal	24	15.7 (8.2–26.9)		
<i>Hemiplegics</i>				
Type I	8	74.9 (36.5–187.6)	51.0 (28.3–91.14)	98.8 (28.5–311.6)
Type II	8	133.32 (32.1–357.0)	103.37 (26.0–276.9)	163.3 (38.2–437.1)
Type III	10	172.4 (55.3–362.0)	70.4 (32.1–94.2)	274.5 (60.70–638.9)
Type IV	4	639.3 (212.2–1066.5)	378.7 (140.2–617.3)	900.0 (284.3–1515.7)
All	30	166.7 (32.1–1066.5)	95.4 (26.0–617.3)	236.0 (28.5–1515.7)
Diplegic	23	279.4 (28.46–1322.3)		
Quadriplegic	11	491.0 (121.5–1195.0)		

^a Right and left side indices averaged to obtain single index for each subject.

hemiplegia; higher for Type IV than Type III; higher for diplegic patients than hemiplegic patients; higher for quadriplegic patients than for diplegic patients; and higher on the affected side of the hemiplegic patients than on the unaffected side. Table 2 indicates that this is, in fact, the case.

The range of values for each diagnosis category was found to be large (Table 2). This is consistent with the wide range of patient characteristics and the imprecise

of the diagnosis categories. However, the mean index values were found to increase with the level of involvement. The mean index for the control subjects was the lowest. The mean index for quadriplegic patients the highest. No overlap was found between index values for the control subjects and the individuals with gait abnormalities. Within the hemiplegic patients the index was seen to increase with type. Type I hemiplegics on average having a lower index than the type II

hemiplegics, type II having a lower index than type III, and type III lower than type IV. The indexes calculated for the affected side for the hemiplegic patients were also seen on average to be larger than the index for the unaffected side. In addition, the range of index values for the affected side was seen to be greater than the range for the unaffected side. Because of the wide range of index values in each of the diagnosis categories, however, no statistical conclusions about the relationship between diagnosis and index values could be made.

4. Discussion

The purpose of the normalcy index is to find a single number that reflects the amount by which a subject's gait deviates from an average normal gait. By default, the notion of normal is subjective. The method proposed, however, is general in nature and allows for subject group or task specific choices of the defining variables. For this study, 16 kinematic parameters were selected. The selection was driven largely by clinical input, but partly by convenience. The 16 variables presented are the 'best effort' of the authors. One should not conclude, however, that other 'better' sets of variables could not be found. The list presented here included only kinematic variables. The inclusion of kinetic variables, while useful in planning intervention, would have excluded non-independent ambulators from the analysis. This study focuses on the task of level ground walking. If the activity of interest were stair climbing, rising from a chair or jumping a completely different set of variables would likely be warranted.

The normalcy index shows potential as a useful tool to objectively quantify overall changes in gait. The index accounts for the correlation between many inter-related gait variables and provides a single number that indicates how closely a given gait pattern approaches normal. As such this index may prove to be useful for outcomes studies in which a means of accurately quantifying gait changes is needed, and for other studies in which there is a need to account for the relationship between correlated gait variables.

The relationship seen between the expected and actual distribution among the subjects with no gait abnormalities and the trends seen between the severity of the diagnosis and the normalcy index gives added confidence that this is a useful measure. While these findings should not be construed as proof of the validity of the measure, in both cases the results compared favorably with expectations. These findings lend support to the notion that the index provides useful information about the nature of the subject's gait.

The value chosen to reflect normalcy (d) is only one of a variety of possible measures that could be applied to a selected set of gait variables. The Euclidean length

based measure has the advantage of being symmetric, invariant under a rotation of the coordinate axes, obeying the triangle inequality and being familiar to most readers [12]. However, it is important to stress that there is no de facto reason to select this distance measure over other possibilities. If, for example, the largest eigenvalue of the covariance matrix dominates the subsequent eigenvalues, its value alone could be interpreted as suitably summarizing the variation inherent in all of the variables. Furthermore, if large deviations from normal are to be severely penalized powers > 2 could be used in Eq. (7). These points highlight the fact that the normalcy index contains inherent subjectivity and must be applied in conjunction with appropriate clinical interpretation and evaluation.

It is worth noting that several other multivariate statistical methods such as factor analysis, cluster analysis, linear and non-linear discrimination analysis and regression analysis are related to the procedure described here. Factor analysis attempts to enumerate a minimal set of underlying factors that influence the observed variables. Cluster analysis is useful for finding similarities and differences in measured variables for subjects drawn from a discrete number of homogeneous groups (such as hemiplegia types). Discrimination analysis is complementary to cluster analysis and is useful for categorizing subjects into one of several known groups using combinations of measured variables. Regression analysis looks for causal relationships whereby dependent variables are determined by independent variables. Principal component analysis, as used in this study, is a method of transforming variables in order to remove their interdependence. In this regard, principal component analysis may be used as a pre-cursor to any of the above mentioned multivariate techniques. For example, a discriminating function may be constructed for categorizing subjects based on their normalcy index along with other variables.

The important clinical question following any treatment protocol is 'Has the patient improved?' From the patient's point of view, the criteria for judging improvement is typically whether his/her function has improved. Is the individual able to do the things he/she wants to do better than before the surgical procedure? Even though the normalcy index presented here does provide a way to quantify and evaluate changes in gait patterns resulting from treatments, it does not necessarily indicate whether there has been any improvement or degradation of function. There is no established relationship between normalcy of gait pattern and function, since abnormalities in gait pattern can actually represent very functional coping mechanisms [16]. However, the reality is that clinical diagnosis is typically made by comparing a patient's measurements to normal data.

Clinical gait analysis laboratories are no exception, commonly plotting a patient's gait data on the same

page with normal controls for comparison. The normalcy index proposed here provides the same type of comparison condensed to a single number. In order to provide a more complete appraisal of a particular patient's improvement the normalcy index will have to be used in conjunction with other outcome measures such as patient satisfaction and/or an evaluation tool for assessing functional improvement. In addition, the particular variables chosen for this paper were based on the empiric experience of the authors. Consequently, it could certainly be argued that another set of variables exists within the data that would more accurately reflect the patient's outcome, for example, a data set that includes selected kinetic variables.

5. Conclusion

A new method has been described for measuring the effects that specific treatments have on a patient's gait. Initial efforts to gain confidence in this measure indicate that it has clinical relevance and is robust. The method has potential as a tool for quantifying surgical outcomes and addressing the question of whether, in the presence of post-operative changes in ambulation, an individual's gait has actually improved. Opportunities for further research still remain. Among these are optimization of the variable set, investigation of non-Euclidean distance based measures, and the possible relationships between the normalcy index and other multivariate statistical methods.

References

- [1] Davis RB, Ounpuu S, Tyburski D, Gage JR. A gait analysis data collection and reduction technique. *Hum Mov Sci* 1991;10:575–87.
- [2] Boscarino LF, Ounpuu S, Davis RB, Gage JR, DeLuca PA. Effects of selective dorsal rhizotomy on gait in children with cerebral palsy. *J Pediatr Orthop* 1993;13:174–9.
- [3] Ounpuu S, Muik E, Davis RB, Gage JR, DeLuca PA. Rectus femoris surgery in children with cerebral palsy. Part I: the effect of rectus femoris transfer location on knee motion. *J Pediatr Orthop* 1993;13:325–30.
- [4] Ounpuu S, Muik E, Davis RB, Gage JR, DeLuca PA. Rectus femoris surgery in children with cerebral palsy. Part II: A comparison between the effect of transfer and release of the distal rectus femoris on knee motion. *J Pediatr Orthop* 1993;13:331–5.
- [5] Rose SA, DeLuca PA, Davis RB, Ounpuu S, Gage JR. Kinematic and kinetic evaluation of the ankle after lengthening of the gastrocnemius fascia in children with cerebral palsy. *J Pediatr Orthop* 1993;13:727–32.
- [6] Winters TF, Gage JR, Hicks R. Gait patterns in spastic hemiplegia in children and young adults. *J Bone Joint Surg* 1987;69A:437–41.
- [7] Winter DA. *The Biomechanics and Motor Control of Human Gait: Normal Elderly and Pathological*. Waterloo, Ont., Canada: University of Waterloo Press, 1991.
- [8] Mah CD, Hulliger M, Lee RG, Callaghan IS. Quantitative analysis of human movement synergies: constructive pattern analysis for gait. *J Motor Behav* 1994;26(2):83–102.
- [9] Laassel EM, Loslever P, Angue JC. Patterns of relations between lower limb angle excursions during normal gait. *J Biomed Eng* 1992;14(4):313–20.
- [10] Loslever P. Error and data coding in the multi-dimensional analysis of human movement signals. *J Eng Med* 1993;207(2):103–7.
- [11] Loslever P, Laassel EM, Angue JC. Combined statistical study of joint angles and ground reaction forces using component and multiple correspondence analysis. *IEEE Trans Biomed Eng* 1994;41(12):1160–7.
- [12] Kendall M. *Multivariate Analysis*. London: Charles Griffin, 1975.
- [13] Kadaba MP, Ramakrishnan HK, Wooten ME. Measurement of lower extremity kinematics during level walking. *J Orthop Res* 1990;8(3):383–92.
- [14] Beyer WH. *CRC Standard Mathematical Tables*, 28th edition. Boca Raton, FL: CRC Press, 1987.
- [15] Box GEP, Hunter WG, Hunter JS. *Statistics for Experimenters: An Introduction to Design, Data Analysis and Model Building*. New York: John Wiley and Sons, 1978.
- [16] Gage JR. *Gait Analysis in Cerebral Palsy*. Clinics in Developmental Medicine, vol. 121. Oxford: MacKeith Press, 1991.