data analysis
IN THE
LIFE SCIENCES

PROCEEDINGS OF A WORKSHOP

Feb. 1984

Edited by

T. KRISHNAN

INDIAN STATISTICAL INSTITUTE
CALCUTTA
INDIA

Chapter Two

FUNCTIONAL ASSESSMENT OF KNEE AND ANKLE DURING LEVEL WALKING

N. Bhadra*, D. Chaturvedi, T.S. Kutty, P.P. Majumder**, G. Poduval and J.S. Rao

1. Introduction

A large part of orthopaedic practice has to deal with patients who have disorders resulting in an abnormal gait. Medical training teaches the art of trying to infer the cause of the abnormality from visual observation of walking. The walking pattern of an individual is distinctive and innumerable objective features may be used to characterise gait. Of the various kinematic events which occur during walking, sagittal plane angular movements (flexion-extension) at the knee and ankle are extremely important in determining normal walking patterns (see e.g., Saunders et al., 1953; Lamoreux, 1971; Bhadra and Bose, 1984).

Objectives :

The aim of this study, albeit preliminary, is to :

(i) find out how well the quantitative angular measurements on flexion-extension of the knees and ankles can be used to characterise the various patient groups identified through clinico-pathological observations;

(ii) find out which of the flexion-extension variables are particularly useful to identify an individual as belonging to one of various patient groups, and, in particular, to identify whether an individual is different from a control individual ;

---

* scientist who supplied problem and data set

** coordinator of data analysis

(iii) obtain a measure to indicate the similarity of an individual with a control individual;

(iv) obtain suggestions for a similar study to be conducted on a large scale.

2. Experimental Method

Knee and ankle goniometers were fabricated using precision, linear carbon-film potentiometers. They were attached with Velcro bands, being centred over the joint axis using the bony landmarks. Two flat foot switches below the feet were used to record the foot events.

Measurements were made with the subjects walking at their own pace along a 8-metre-long walkway turning around at each end. Records were made on a four-track chart recorder (POLYRITE) from one side of the subject at a time.

Four traverses were recorded with the electrogoniometers fixed on one side. This procedure was preceded by a clinical and anthropometric examination.

3. The Data

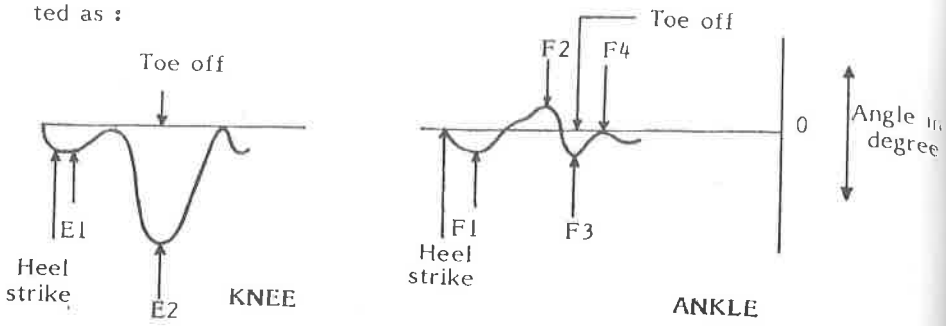
The sample : A total of thirtynine individuals were selected for this study. These individuals are drawn from eight groups - seven of these groups comprise individuals who have, in some form or the other, unilateral orthopaedic impairment, and the eighth group comprises individuals who are normal in the sense that they have no orthopaedic impairment. The various groups, the codes by which these are referred to (in the later sections), brief descriptions of the groups and the sample sizes are given below :

Sl. No.	Group Name	Code	Description	Sample Size
1.	Control	CO	Subjects with no locomotor complaints and with no signs of locomotor disorder on physical examination	10
2.	Degenerative Arthritis	DA	Subjects with unilateral osteoarthrosis of the knee	5
3.	Rheumatoid Arthritis	RA	Subjects with unilateral knee disorder diagnosed as rheumatoid arthritis	5
4.	Post-Traumatic Stiffness	PT	Subjects with stiffness of knee or ankle or both of one side following direct or indirect trauma	5
5.	Post-Menisectomy	PM	Subjects who underwent operative removal of the medial meniscus in one knee	4
6.	Crutch-User	CR	Subjects with one unaffected lower limb using axillary crutches for three-point crutch walking	5
7.	Below-knee Amputee	BK	Unilateral below-knee amputees using Patellar-Tendon Bearing prostheses with solid-ankle-cushioned heel footpieces	3
8.	Through-knee Amputee	TK	Unilateral through-knee amputees using a quadrilateral socket single axis external knee-hinge prostheses with SACH foot pieces	2
	TOTAL			39

The Variables :

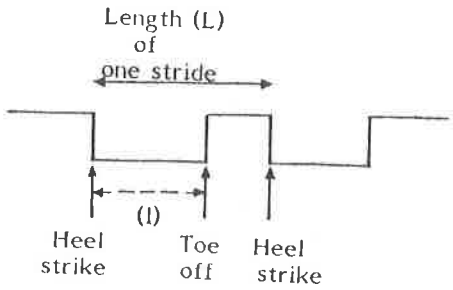
During bipedal walking, the feet are lifted off and placed alternately on the ground, with intervening periods when both feet are on the ground. A single walking cycle of one limb consists of a period when the foot

is on the ground (stance) followed by the next phase of forward movement (swing). A stride is the total event of one limb from the start of one stance phase to the next stance phase. The sequential events are heel strike, foot flat, toe strike, heel off, toe off and then the swing phase upto the next heel strike. Thus pictorially a stride can be depicted as :

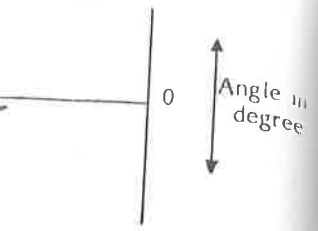


The ratio $\frac{l}{L}$ denotes the length of time the foot is on the ground relative to the total time period of one stride; $100\ \frac{l}{L}$ is called the percent Stance (ST).

During a stride the knee and the ankle undergoes flexion and extension, which are angular movements, and can be recorded graphically from the electrogoniometers. These are usually of the following types:



phase of forward move-
one limb from the start
The sequential events
e off and then the swing
y a stride can be depic-



ANKLE

is on the ground rela-
is called the per-

goes flexion and exten-
e recorded graphically
of the following types:

The Peak Excursion Variables (E1, F1, etc.) are called :

E1 = Knee Stance Flexion,
E2 = Knee Swing Flexion,
F1 = Ankle Stance Flexion,
F2 = Ankle Stance Extension,
F3 = Ankle Swing Flexion,
F4 = Ankle Swing Extension.

Thus, for each side of a sampled individual, observations are obtained on the seven variables defined above (E1, E2, F1,...,F4, ST). Since, normally, each individual has two legs, there are 14 observations per individual.

The Observations Available :

The observations on the variables for each individual are the mean values based on 20 strides - the raw data were not available. Also, because of orthopaedic impairment, most sampled individuals have one affected leg. We shall, therefore, denote the variables pertaining to both the legs as : E1A, E2A, ..., F4A, STA, for the affected leg, and E1U, E2U, F1U, ..., F4U, STU, for the unaffected leg. Since the control individuals are unaffected in both legs, we shall use the suffixes A and U to denote, respectively, the right and left legs. Moreover, because of the nature of the orthopaedic impairment, the CR group individuals do not have any observations on variables E1A, E2A, F1A, ..., F4A, STA, and the individuals in groups BK and TK do not have any observations on variables F1A, ..., F4A. Thus, there are 7 variables common to all the 8 groups, 10 variables common to the 7 groups excluding CR, and 14 variables common to the 5 groups CO, DA, RA, PT, and PM.

Limitations of the Data :

Since the data for each individual consist of only mean values, and not the raw data for each of the 20 strides based on which the mean values were calculated, the intra-individual variability could not

be measured. Second, the number of individuals in each group is very small. This seriously affected the testing of various assumptions and hypotheses concerning parameters of the statistical models. In most cases, we were, therefore, forced to take for granted that the various assumptions underlying the statistical methods and models used were satisfied by the data. Third, the sample size limitation did also not permit us to use important covariates, such as age, sex, etc., and eliminate their effects prior to subjecting the data to statistical analyses.

4. The Complete Data Set :

Data on Functional Assessment of knee and ankle during level walking

Source : Bio-Engineering Unit, Department of Orthopaedics, University College of Medicine, Calcutta University.

The data comprise angular movement variables of knee and ankle during level walking, for each leg, these variables are :

E1 = Knee Stance Flexion,
E2 = Knee Swing Flexion,
F1 = Ankle Stance Flexion,
F2 = Ankle Stance Extension,
F3 = Ankle Swing Flexion,
F4 = Ankle Swing Extension,
ST = Percent Stance.

The data are the mean values of actual observations based on 20 strides. Each sampled individual comes from one of 8 groups, the groups being :

1 = Control,
2 = Degenerative Arthritis,
3 = Rheumatoid Arthritis,
4 = Post-Traumatic Stiffness,
5 = Post-Menisectomy,
6 = Crutch-User,
7 = Below-knee Amputee,
8 = Through-Knee Amputee.

Obviously for members of group 1, both legs are normal, whereas the individuals belonging to the 7 groups being unilateraliy orthopaedically handicapped, have one leg affected (the other leg normal). Note also that certain variables are unobservable in individuals belonging to some groups because of loss of one leg.

The complete data set is presented in Table 1.

5. A Comment on Statistical Treatment of the Data

Since all the Peak Excursion values (Variables : E1, E2, F1,...,F4) are in degrees, the use of standard statistical methods for linear data (Rao, 1952) in analysing the present angular data may be questioned. However, although the present data are angular, the range of variation for each variable is very small, within 20°. Since the range of variation is within one quadrant of a circle, linear treatment of the present angular data does not create any problem (Mardia et al., 1979). We have, in fact, checked that the mean values of each variable do agree fairly well when the data are treated as linear data in comparison with the data treated as angular data. We have also performed the analyses of variance for a couple of variables, treating the data as angular. It was seen that the results were in very good agreement with those obtained upon a linear treatment of the data.

6. Descriptive Statistics

The means, standard deviations and the ranges of the variables separately for each of the eight groups are given in Tables 2(A) - (C). It may be recalled that for some of the groups, because of the nature of orthopaedic impairment, observations on certain variables are missing. From these tables, it is seen that for most of the orthopaedically impaired groups there is a general tendency of reduction in the total amount of flexion/extension (as is evident from both the mean values and the ranges of the Peak Excursion Variables for the Knee and Ankle) in comparison with the control individuals. For several variables, the extent of within-group variability seems to be different for the different groups. However, we are forced to ignore these differences because of the small sample sizes involved. We note that there is no consistent correspondence among mean values and standard deviations of the variables, and hence we have not considered any transformation of the data. Whether the data can be assumed to have come from Normal distributions have, however, not been tested because of the small sample sizes involved.

Table 1 : The complete data set

Group	Sl. No.	Right/Unaffected Leg							Left/Affected Leg						
		E1U	E2U	F1U	F2U	F3U	F4U	STU	E1A	E2A	F1A	F2A	F3A	F4A	STA
1	1	19	71	-9	16	-13	4	59.6	19	71	-9	10	-15	5	59.6
1	2	17	66	-10	15	-14	2	60.0	17	65	-10	15	-14	2	60.0
1	3	22	66	-13	7	-15	2	59.3	19	68	-8	8	-14	2	59.3
1	4	18	63	-11	6	-15	0	59.3	18	65	-10	6	-14	0	59.3
1	5	16	70	-12	10	-14	0	57.2	15	68	-12	12	-11	0	57.2
1	6	25	64	-12	11	-10	0	61.1	26	65	-10	10	-11	0	61.1
1	7	16	64	-11	13	-10	0	59.2	14	64	-11	10	-9	0	59.2
1	8	15	59	-10	12	-10	1	59.4	14	60	-10	11	-10	1	59.4
1	9	17	58	-12	15	-9	0	59.3	21	57	-14	14	-10	0	59.3
1	10	16	59	-10	10	-12	3	60.0	15	63	-9	18	-13	3	60.0
2	1	15	62	-18	18	-13	1	68.2	6	55	-20	12	-11	1	54.2
2	2	14	72	-16	10	-20	-6	68.4	6	44	-14	12	-20	-8	63.1
2	3	12	54	-10	14	-12	0	59.6	9	36	-8	8	-16	4	52.2
2	4	15	66	-10	10	-18	0	67.7	10	52	-10	10	-12	-4	62.1
2	5	14	60	-12	12	-15	0	63.2	6	42	-10	12	-12	0	62.4
3	1	18	72	-16	14	-22	-6	62.3	9	35	-20	14	-10	0	61.1
3	2	10	64	-16	14	-18	-4	68.4	7	35	-14	12	-18	-2	60.4
3	3	15	69	-6	15	-18	0	63.2	10	38	-12	21	-21	-3	60.2
3	4	12	69	-12	16	-15	0	59.1	10	42	-6	21	-24	-6	58.3
3	5	18	69	-18	9	-24	3	68.2	7	36	-16	12	-18	3	60.3
4	1	14	44	-15	18	-12	0	68.2	20	64	-15	15	-3	3	61.1
4	2	12	48	-18	9	-15	0	71.1	15	54	-12	9	-3	0	62.2
4	3	10	50	-4	16	-10	0	69.0	14	68	-8	10	-4	0	61.1
4	4	12	60	-10	14	-12	0	68.2	15	69	-6	10	-4	0	60.4
4	5	10	50	-4	18	-20	0	66.7	0	34	-4	24	-10	0	60.0
5	1	15	69	-12	15	-12	0	62.1	12	63	-9	15	-15	0	60.0
5	2	18	64	-12	18	-15	0	60.0	15	60	-12	18	-10	0	54.2
5	3	16	64	-12	14	-10	0	61.3	14	62	-12	16	-14	2	60.2
5	4	14	66	-12	16	-10	0	66.2	12	60	-12	16	-12	0	61.1
6	1	12	44	-12	6	-6	0	67.6	999	999	999	999	999	999	99.9
6	2	16	44	-9	15	-12	0	69.0	999	999	999	999	999	999	99.9
6	3	15	45	-12	9	-6	0	66.6	999	999	999	999	999	999	99.9
6	4	12	69	-6	9	-12	0	60.0	999	999	999	999	999	999	99.9
6	5	15	60	-6	15	-12	0	62.8	999	999	999	999	999	999	99.9
7	1	24	66	-12	18	-14	-2	62.5	5	50	999	999	999	999	45.2
7	2	12	60	-12	15	-21	0	67.5	16	48	999	999	999	999	55.0
7	3	22	62	-10	16	-16	0	66.7	12	48	999	999	999	999	50.1
8	1	20	64	-10	20	-20	2	77.1	0	26	999	999	999	999	62.9
8	2	20	72	-6	14	-18	0	75.6	0	20	999	999	999	999	62.2

NOTE : '99.9' and '999' indicate unobservable data.

Table 2(A) : Descriptive Statistics for Peak Excursion Variables for the Knees

Variable	Group	Statistic				
		Mean ± s.e.	s.d.	Max.	Min.	Range
E1A	CO	17.8±1.18	3.74	26	14	12
	DA	7.4±0.87	1.95	10	6	4
	RA	8.6±0.68	1.52	10	7	3
	PT	12.8±3.37	7.53	20	0	20
	PM	13.2±0.75	1.50	15	12	3
	CR	-	-	-	-	-
	BK	11.0±3.22	5.57	16	5	11
	TK	0.0±0.00	0.00	0	0	0
E2A	CO	64.6±1.28	4.03	71	57	14
	DA	45.8±3.44	7.69	55	36	19
	RA	38.0±1.22	2.74	42	35	7
	PT	57.8±6.52	14.57	69	34	35
	PM	61.2±0.75	1.50	63	60	3
	CR	-	-	-	-	-
	BK	48.7±0.67	1.16	50	48	2
	TK	23.0±3.00	4.24	26	20	6
E1U	CO	18.1±0.99	3.14	25	15	10
	DA	14.0±0.55	1.22	15	12	3
	RA	14.6±1.60	3.58	18	10	8
	PT	11.6±0.75	1.67	14	10	4
	PM	15.8±0.85	1.71	18	14	4
	CR	14.0±0.84	1.87	16	12	4
	BK	19.3±3.71	6.83	24	12	12
	TK	20.0±0.00	0.00	20	20	0
E2U	CO	64.0±1.41	4.47	71	58	13
	DA	62.8±3.01	6.72	72	54	18
	RA	68.6±1.29	2.88	72	64	8
	PT	50.4±2.64	5.90	60	44	16
	PM	65.8±1.18	2.36	69	64	5
	CR	52.4±5.14	11.50	69	44	25
	BK	62.7±1.76	3.06	66	60	6
	TK	68.0±4.00	5.66	72	64	8

Table 2(B.1): Descriptive Statistics for Peak Excursion Variables for the Ankles (Unaffected leg)

Variable	Group	Statistic				
		Mean ± s.e.	s.d.	Max.	Min.	Range
F1U	CO	-11.0±0.39	1.25	-9	-13	4
	DA	-13.2±1.62	3.63	-10	-18	8
	RA	-13.6±2.14	4.78	-6	-18	12
	PT	-10.2±2.84	6.34	-4	-18	14
	PM	-12.0±0.00	0.00	-12	-12	0
	CR	- 9.0±1.34	3.00	- 6	-12	6
	BK	-11.3±0.67	1.16	-10	-12	2
	TK	- 8.0±2.00	2.83	- 6	-10	4
F2U	CO	12.3±1.23	3.89	18	6	12
	DA	12.8±1.50	3.35	18	10	8
	RA	13.6±1.21	2.70	16	9	7
	PT	15.0±1.67	3.74	18	9	9
	PM	15.8±0.85	1.71	18	14	4
	CR	10.8±1.80	4.02	15	6	9
	BK	16.3±0.88	1.53	18	15	3
	TK	17.0±3.00	4.24	20	14	6
F3U	CO	-12.2±0.73	2.30	-9	-15	6
	DA	-15.6±1.50	3.36	-12	-20	8
	RA	-19.4±1.60	3.58	-15	-24	9
	PT	-13.8±1.74	3.90	-10	-20	10
	PM	-11.8±1.18	2.36	-10	-15	5
	CR	- 9.6±1.47	3.27	- 6	-12	6
	BK	-17.0±2.08	3.61	-14	-21	7
	TK	-19.0±1.00	1.41	-18	-20	2
F4U	CO	1.2±0.47	1.48	4	0	4
	DA	- 1.0±1.26	2.83	1	- 6	7
	RA	- 1.4±1.60	3.58	3	- 6	9
	PT	0.0±0.00	0.00	0	0	0
	PM	0.0±0.00	0.00	0	0	0
	CR	0.0±0.00	0.00	0	0	0
	BK	- 0.7±0.67	1.16	0	- 2	2
	TK	1.0±1.00	1.41	2	0	2

k Excursion Variables
)

Min.	Range
-13	4
-18	8
-18	12
-18	14
-12	0
-12	6
-12	2
-10	4
6	12
10	8
9	7
9	9
14	4
6	9
15	3
14	6
-15	6
-20	8
-24	9
-20	10
-15	5
-12	6
-21	7
20	2
0	4
6	7
6	9
0	0
0	0
0	0
2	2
0	2

Table 2(B.2) : Descriptive Statistics for Peak Excursion Variables for the Ankles (Affected leg)

Variable Group		Statistic				
		Mean ± s.e.	s.d.	Max.	Min.	Range
F1A	CO	-10.3±0.54	1.70	-8	-14	6
	DA	-12.4±2.14	4.78	-8	-20	12
	RA	-13.6±2.32	5.18	-6	-20	14
	PT	- 9.0±2.00	4.47	-4	-15	11
	PM	-11.2±0.80	1.50	-9	-12	3
	CR	-	-	-	-	-
	BK	-	-	-	-	-
	TK	-	-	-	-	-
F2A	CO	11.4±1.11	3.50	18	6	12
	DA	10.8±0.80	1.79	12	8	4
	RA	16.0±2.07	4.64	21	12	9
	PT	13.6±2.80	6.27	24	9	15
	PM	16.2±0.63	1.26	18	15	3
	CR	-	-	-	-	-
	BK	-	-	-	-	-
	TK	-	-	-	-	-
F3A	CO	-12.1±0.67	2.13	-9	-15	6
	DA	-14.2±1.68	3.77	-11	-20	9
	RA	-18.2±2.33	5.22	-10	-24	14
	PT	- 4.8±1.32	2.95	- 3	-10	7
	PM	-12.8±1.11	2.22	-10	-15	5
	CR	-	-	-	-	-
	BK	-	-	-	-	-
	TK	-	-	-	-	-
F4A	CO	1.3±0.54	1.70	5	0	5
	DA	- 1.4±2.09	4.67	4	- 8	12
	RA	- 1.6±1.50	3.36	3	- 6	9
	PT	0.6±0.60	1.34	3	0	3
	PM	0.5±0.50	1.00	2	0	2
	CR	-	-	-	-	-
	BK	-	-	-	-	-
	TK	-	-	-	-	-

Table 2(C) : Descriptive Statistics for Stance Variables

Variable Group		Statistic				
		Mean ± s.e.	s.d.	Max.	Min.	Range
STA	CO	59.44±0.31	0.97	61.1	57.2	3.9
	DA	58.80±2.31	5.17	63.1	52.2	10.9
	RA	60.06±0.47	1.04	61.1	58.3	2.8
	PT	60.96±0.38	0.84	62.2	60.0	2.2
	PM	58.88±1.58	3.15	61.1	54.2	6.9
	CR	-	-	-	-	-
	BK	50.10±2.83	4.90	55.0	45.2	9.8
	TK	62.55±0.35	0.50	62.9	62.2	0.7
STU	CO	59.44±0.31	0.97	61.1	57.2	3.9
	DA	65.42±1.74	3.89	68.4	59.6	8.8
	RA	66.24±1.44	3.22	69.1	62.3	6.8
	PT	68.64±0.72	1.61	71.1	66.7	4.4
	PM	62.40±1.34	2.68	66.2	60.0	6.2
	CR	65.54±1.79	4.00	69.7	60.0	9.7
	BK	65.57±1.55	2.69	67.5	62.5	5.0
	TK	76.35±0.75	1.06	77.1	75.6	1.5

ce Variables

Min.	Range
57.2	3.9
52.2	10.9
58.3	2.8
60.0	2.2
54.2	6.9
-	-
45.2	9.8
62.2	0.7
57.2	3.9
59.6	8.8
62.3	6.8
66.7	4.4
60.0	6.2
60.0	9.7
62.5	5.0
75.6	1.5

For purposes of illustration, we consider the three variables E1A, E1U, E2A. If data on these three variables are treated as angular, then the mean values in five of the groups are as follows :

Group	Mean		
	E1A	E1U	E2A
CO	17.8	17.9	64.6
DA	7.4	13.9	45.8
RA	8.6	14.6	38.0
PT	12.8	11.6	57.9
PM	13.2	15.8	61.2

These values obviously correspond very well with the values presented in Table 2(A).

7. A Note on the Statistical Analysis of the Data

As we have mentioned earlier, because of the nature of orthopaedic impairments among individuals belonging to certain groups, observations on some of the variables could not be obtained. We have, therefore, treated the data as three separate sets as described below in order to make the analyses as complete as possible :

Data Set 1 : Groups (5) = (CO, DA, RA, PT, PM)

Variables (14) = (E1A, E2A, E1U, E2U, F1A, ... , F4A, F1U, ... , F4U, STA, STU)

Total sample size = 29.

Data Set 2 : Groups (7) = (CO, DA, RA, PT, PM, BK, TK)

Variables (10) = (E1A, E2A, E1U, E2U, F1U, F2U, F3U, F4U, STA, STU)

Total sample size = 34.

Data Set 3 : Groups (8) = (CO, DA, RA, PT, PM, CR, BK, TK)

Variables (7) = (E1U, F1U, F2U, F3U, F4U, STU) : all pertaining to the unaffected leg.

Total sample size = 39.

8. Analysis of Data Set 1

The first point that we wanted to verify was whether individuals belonging to the same group were more similar than individuals belonging to two different groups. This had to be true at least for some of the groups, or else the exercise of trying to identify variables for discriminating between the groups and of trying to find whether an individual belonging to a particular orthopaedically handicapped group is closer to a control individual than another individual would be futile. For this, we used the Euclidean Distance based on all the 14 variables and the Single Linkage method to cluster the 29 individuals belonging to the 5 groups. The results are given in the form of a dendrogram in Figure 1. From this figure it is seen that individuals belonging to a particular group, do, by and large, cluster together.

Then, in order to identify variables which show significant differences in mean values among the various groups, we performed analyses of variance. The results are given in Table 3, from which it is seen that only seven (E1A, E1U, E2A, E2U, F3A, F3U, STU) of the fourteen variables show significant differences in mean values among the groups. In order to find out how the findings obtained upon linear treatment of the data compare with those obtained by treating the data as angular, we performed an analysis of variance for a couple of variables (E2A and F1A) treating these as angular. The results correspond very well with those obtained upon linear treatment of the data : the new F-ratios with 4 and 24 d.f. are 14.1125 and 1.3, respectively, for variables E2A and F1A .

, CR, BK, TK)

STU) : all pertaining

was whether individuals
than individuals belong-
ie at least for some of
ntify variables for dis-
nd whether an individual
apped group is closer
would be futile. For
ll the 14 variables and
ndividuals belonging to
m of a dendrogram in
viduals belonging to a

show significant differ-
we performed analyses
from which it is seen
STU) of the fourteen
ues among the groups.
upon linear treatment
g the data as angular,
ole of variables (E2A
correspond very well
ata : the new F-ratios
ly, for variables E2A

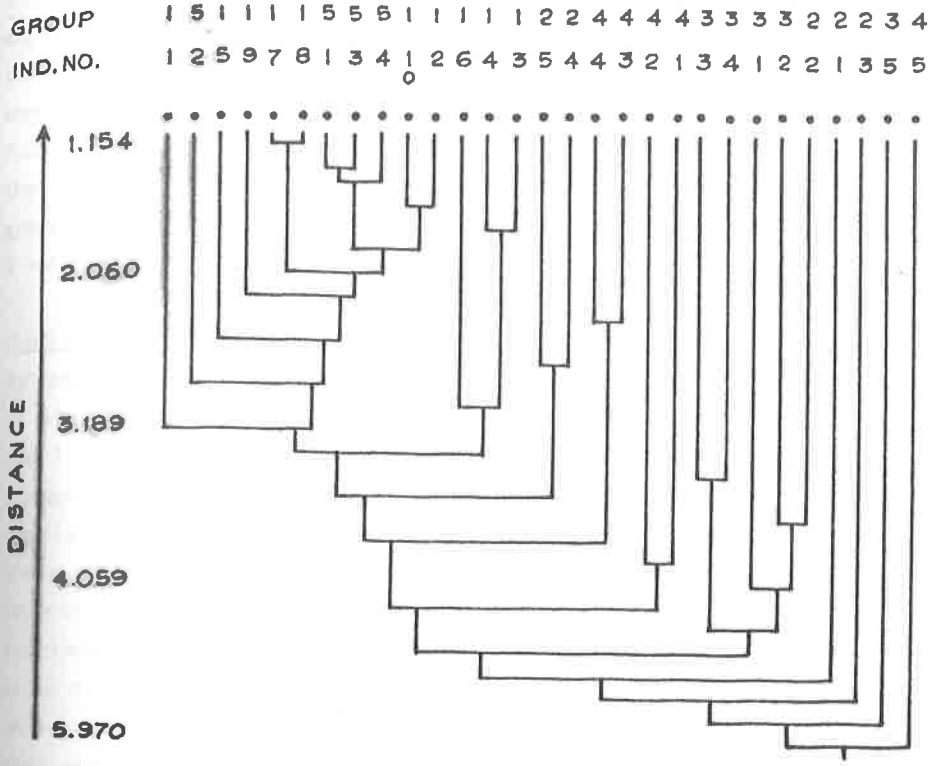


Figure 1 : Dendrogram showing relationships among individuals belonging to 5 groups (Group 1 = CO, 2 = DA, 3 = RA, 4 = PT, 5 = PM). [Based on 14 variables.]

Table 3 : Test of equality of mean values among five groups (CO, DA,RA,PT,PM) for each of the 14 variables

Variable	Between Groups		Within Groups		F-Value
	M.S.	d.f.	M.S.	d.f.	
E1A	121.0778	4	15.9812	24	7.58*
E2A	746.1743	4	52.8646	24	14.11*
E1U	39.6254	4	6.9187	24	5.73*
E2U	245.8228	4	22.9146	24	10.73*
F1A	17.0495	4	12.9687	24	1.31
F2A	34.5745	4	15.4646	24	2.24
F3A	118.6252	4	10.6687	24	11.12*
F4A	10.4974	4	7.0292	24	1.49
F1U	11.3345	4	13.2833	24	0.85
F2U	12.0806	4	11.4521	24	1.05
F3U	52.4400	4	9.2312	24	5.68*
F4U	7.3000	4	4.2833	24	1.70
STA	3.9630	4	6.3590	24	0.62
STU	89.2980	4	5.9348	24	15.05*

* significant at the 5 per cent level

ong five groups (CO,
riables

ps	F-Value
.f.	
24	7.58*
24	14.11*
24	5.73*
24	10.73*
24	1.31
24	2.24
24	11.12*
24	1.49
24	0.85
4	1.05
4	5.68*
4	1.70
4	0.62
4	15.05*

Further, since it is also of interest to find for which of the variables the various orthopaedically impaired groups differ from the control group, we performed t-tests for testing equality of mean values between the control group and each of the othopaedically impaired groups for each of the 14 variables. The results are given in Table 4.. From this table the following pictures emerge : (i) the RA group seems to be the most dissimilar to the CO group, as 12 of the 14 variables show significant differences with the CO group; (ii) the PM group seems to be the most similar to the CO group; (iii) all the orthopaedically impaired groups show significant differences with the CO group in respect of 2 variables - STA and STU.

The next question we asked is : are all the 7 variables significantly important for purposes of discriminating between the groups? To answer this, we performed a stepwise discriminant analysis, results of which are presented in Table 5. From this table it is seen that of the 7 variables which show significant differences in mean values among groups, only three (STU, E2A, E2U : in that order of importance) are significant for purposes of discrimination; given the observations on these three variables, the remaining variables do not contribute significantly to the discriminating power. Using the estimated classification functions, we have also classified the individuals into the 5 groups in order to get an idea about the performance of the discriminating variables. These results are also shown in Table 5 , from which it is seen that only 3 of the 29 individuals are misclassified; the probability of correct classification using only 3 of the 14 variables is about 0.9. The classification of individuals was done assuming equal prior probabilities for each of the 5 groups. Based upon the classification functions, the posterior probabilities of group inclusion were also obtained for each of the 29 individuals. These probabilities are presented in Table 6. These probabilities can be used as a measure of similarity of an individual to the various groups. To exemplify, suppose we ask how

Table 4 : Comparison of Orthopaedically Impaired Groups with the Control Group

Groups Compared	E1A	E2A	E1U	E2U	F1A	F2A	F3A	F4A	F1U	F2U	F3U	F4U	STA	STU
CO - DA	5.76*	6.33*	2.77*	0.41	1.28	0.35	1.40	1.67	1.77	-0.24	2.32*	2.01*	3.92*	-47.39*
CO - RA	5.21*	13.19*	1.95*	-2.07*	1.88*	-2.16*	3.28*	2.26*	1.67	-0.66	4.77*	2.03*	-11.41*	-63.34*
CO - PT	1.75	1.42	4.28*	5.01*	-0.83	-0.88	-5.52*	0.80	-0.39	-1.28	1.04	1.78*	-29.78*	-139.54*
CO - PM	2.34*	1.61	1.36	-0.75	0.92	-2.62*	0.54	0.87	1.56	-1.70	-0.29	1.58	5.30*	-31.64*
CO - CR	-	-	2.66*	2.87*	-	-	-	-	-1.86*	0.70	-1.80	1.78*	-	-47.17*
CO - BK	2.50*	6.56*	-0.45	0.46	-	-	-	-	0.36	-1.70	2.82*	2.02*	62.61*	-64.48*
CO - TK	6.48*	13.26*	-0.82	-1.12	-	-	-	-	-2.61*	-1.54	3.94*	0.18	-43.00*	-222.90*

* significant at the 5 per cent level

Table 5 : Results of Stepwise Discriminant Analysis for Data Set 1

(A) CLASSIFICATION FUNCTIONS

Variable	Group				
	CO	DA	RA	PT	PM
STU	9.641	11.176	11.450	11.639	10.264
E2A	-0.402	-0.964	-1.239	-0.590	-0.564
E2U	2.084	2.206	2.590	1.404	2.193
Constant	-341.853	-414.362	-446.136	-419.392	-376.673

(B) CLASSIFICATION MATRIX

Group	% Correctly Classified	No. of Cases Classified into Group					Within Group sl. Nos.of cases Misclassified
		CO	DA	RA	PT	PM	
CO	100.0	10	0	0	0	0	-
DA	80.0	0	4	1	0	0	2
RA	100.0	0	0	5	0	0	-
PT	80.0	0	1	0	4	0	5
PM	75.0	1	0	0	0	3	2
Total	89.7	11	5	6	4	3	-

* significant at the 5 per cent level

Table 6 : Posterior probabilities of group-inclusion for 29 iindividuals belonging to 5 groups

Within-group Individual Sl. No.		Posterior Probabilities for Group				
		CO	DA	RA	PT	PM
CO	1	0.812	0.000	0.000	0.000	0.188
	2	0.687	0.000	0.000	0.000	0.313
	3	0.847	0.000	0.000	0.000	0.153
	4	0.825	0.000	0.000	0.000	0.175
	5	0.930	0.000	0.000	0.000	0.070
	6	0.579	0.000	0.000	0.000	0.421
	7	0.793	0.000	0.000	0.000	0.207
	8	0.752	0.000	0.000	0.000	0.247
	9	0.689	0.001	0.000	0.000	0.310
	10	0.773	0.000	0.000	0.000	0.227
DA	1	0.000	0.889	0.011	0.067	0.033
	2	0.000	0.072	0.927	0.000	0.000
	3	0.005	0.936	0.010	0.001	0.049
	4	0.000	0.882	0.103	0.001	0.015
	5	0.000	0.931	0.049	0.001	0.019
RA	1	0.000	0.034	0.966	0.000	0.000
	2	0.000	0.300	0.700	0.000	0.000
	3	0.000	0.165	0.834	0.000	0.001
	4	0.000	0.106	0.894	0.000	0.000
	5	0.000	0.028	0.972	0.000	0.000
PT	1	0.000	0.000	0.000	1.000	0.000
	2	0.000	0.000	0.000	1.000	0.000
	3	0.000	0.000	0.000	1.000	0.000
	4	0.006	0.012	0.000	0.856	0.126
	5	0.000	0.803	0.021	0.176	0.000
PM	1	0.236	0.004	0.000	0.000	0.761
	2	0.548	0.001	0.000	0.000	0.451
	3	0.427	0.002	0.000	0.000	0.571
	4	0.012	0.375	0.003	0.003	0.606

p-inclusion for 29
ups

or Group	
PT	PM
0.000	0.188
0.000	0.313
0.000	0.153
0.000	0.175
0.000	0.070
0.000	0.421
0.000	0.207
0.000	0.247
0.000	0.310
0.000	0.227
0.067	0.033
0.000	0.000
0.001	0.049
0.001	0.015
0.001	0.019
0.000	0.000
0.000	0.000
0.000	0.001
0.000	0.000
0.000	0.000
.000	0.000
.000	0.000
.000	0.000
.856	0.126
.176	0.000
.000	0.761
.000	0.451
.000	0.571
.003	0.606

similar is individual No.3 of the Post-Meniscus group to a Control individual. From Table 6, we find that the posterior probability of this individual (PM-3) belonging to the CO group is 0.427. Thus, we can say that this individual is about 42.7% similar to a Control individual.

We have thus achieved what we set out to do, viz., to identify variables that afford a maximum discrimination between groups and to obtain a measure of an individual's closeness to the various, in particular the CO, groups.

9. Analysis of Data Sets 2 and 3

Since we have already explained the protocol and the rationale of the various steps and methods for the statistical analysis of the data in the previous section, we shall, in this section, only present the results for data sets 2 and 3 very concisely. The results of Cluster Analysis are presented in the form of dendrograms in Figures 2 and 3 for Data sets 2 and 3, respectively. These figures show that the patterns of clustering of individuals within groups for Data sets 2 and 3 are not as neat as for Data set 1. In fact, for Data set 3, the clustering is fairly poor for most groups. The results of Analyses of Variance are presented in Tables 7 and 8. From these tables, it is seen that there is a very consistent pattern of the variables showing significant between-group differences in mean values for all the three data sets. The results of Stepwise Discriminant Analyses for Data Sets 2 and 3 are presented in Tables 9 and 10, respectively. As is expected from the patterns of clustering, the percentage of correctly classified cases decreases as one moves from Data set 1 to Data set 3. In fact, for Data Set 3 the discriminatory power (as measured by the probability of correct classification) of the two significant discriminatory variables (E2U and STA) is rather low. The interesting point to note is that in none of the 3 data sets do any of the ankle variables enter the discriminant functions, and both knee and stance variables turn out to be significant discriminators for all the 3 data sets. The posterior probabilities of group-inclusion

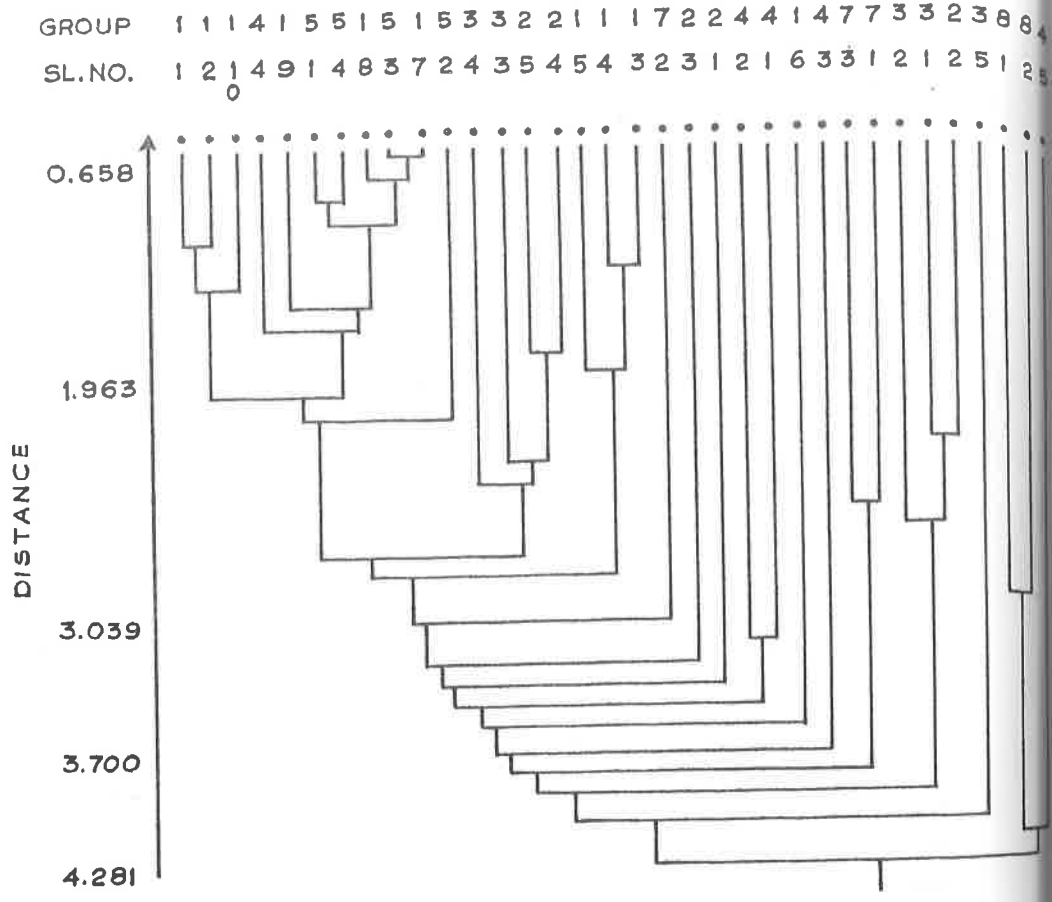
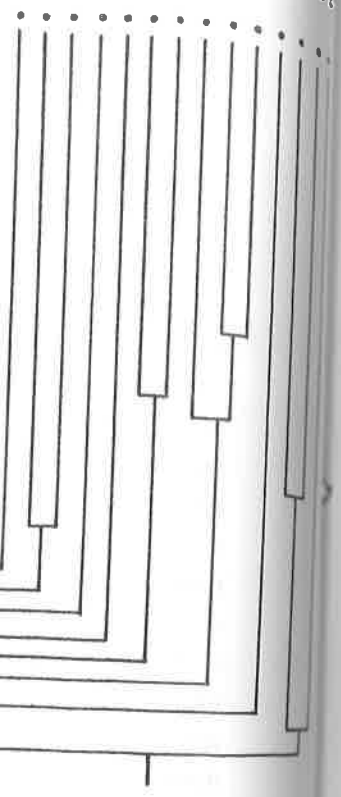


Figure 2 : Dendrogram showing relationships among individuals belonging to 7 groups (Group 1 = CO, 2 = DA, 3 = RA, 4 = PT, 5 = PM, 6 = BK, 7 = TK). [Based on 10 variables.]



lividuals belonging to 7
PT, 5 = PM, 6 = BK,

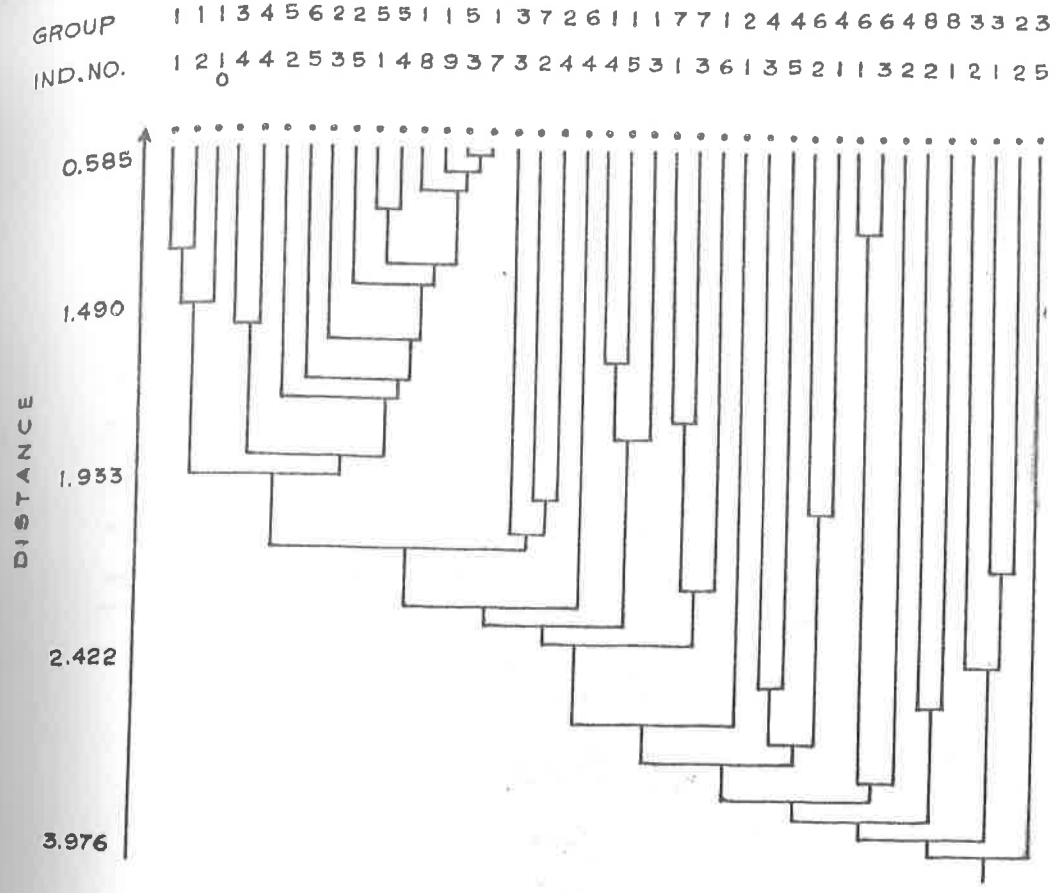


Figure 3 : Dendrogram showing relationships among individuals belonging to 8 groups (Group 1 = CO, 2 = DA, 3 = RA, 4 = PT, 5 = PM, 6 = CR, 7 = BK, 8 = TK). [Based on 7 variables.]

Table 7 : Test of equality of mean values among 7 groups for 10 variables

Variable	Between Groups		Within Groups		F-value
	M.S.	d.f.	M.S.	d.f.	
E1A	133.4082	6	16.5018	27	8.08*
E2A	828.3372	6	47.7562	27	17.35*
E1U	39.3756	6	9.2117	27	4.27*
E2U	173.3873	6	22.2451	27	7.79*
F1U	12.1526	6	12.2025	27	1.00
F2U	14.7472	6	11.0191	27	1.34
F3U	44.7612	6	9.2426	27	4.84*
F4U	5.4222	6	3.9802	27	1.36
STA	48.2973	6	7.4400	27	6.49*
STU	110.4825	6	5.8514	27	18.88*

* significant at the 5 per cent level

es among 7 groups

roups d.f.	F-value
27	8.08*
27	17.35*
27	4.27*
27	7.79*
27	1.00
27	1.34
27	4.84*
27	1.36
27	6.49*
27	18.88*

Table 8 : Test of equality of mean values for 8 groups and 7 variables

Variable	Between Groups		Within Groups		F-value
	M.S.	d.f.	M.S.	d.f.	
E1U	36.1687	7	8.4747	31	4.27*
E2U	216.2751	7	36.4457	31	5.93*
F1U	14.4938	7	11.7892	31	1.23
F2U	19.0170	7	11.6876	31	1.63
F3U	54.9770	7	9.4435	31	5.82*
F4U	4.6476	7	3.4667	31	1.34
STU	95.3024	7	7.1606	31	13.31*

* significant at the 5 per cent level

Table 9 : Results of Stepwise Diiscriminant Analysis for Data Set 2

(A) CLASSIFICATION FUNCTIONS

Variable	Group						
	CO	DA	RA	PT	PM	BK	TK
E2A	-0.024	-0.707	-1.003	-0.280	-0.235	-0.756	-1.668
E2U	2.087	2.361	2.758	1.542	2.272	2.489	3.003
STA	4.162	3.133	3.051	3.595	3.646	1.538	2.191
STU	7.493	9.643	10.017	9.716	8.436	10.558	12.771
Constant	-414.331	-467.431	-500.899	-475.744	-439.948	-446.173	-640.929

(B) CLASSIFICATION MATRIX

Group	% correctly classified	No. of cases classified into groups							Within group sl.nos. of cases misclassified
		CO	DA	RA	PT	PM	BK	TK	
CO	100.0	10	0	0	0	0	0	0	-
DA	60.0	0	3	1	0	0	1	0	BK-1,RA-2
RA	100.0	0	0	5	0	0	0	0	-
PT	80.0	0	1	0	4	0	0	0	5
PM	75.0	1	0	0	0	3	0	0	3
BK	100.0	0	0	0	0	0	3	0	-
TK	100.0	0	0	0	0	0	0	2	-
Total	88.2	11	4	6	4	3	4	2	

PT	80.0	0	1	0	4	0	0	0	5
PM	75.0	1	0	0	0	3	0	0	3
BK	100.0	0	0	0	0	0	3	0	-
TK	100.0	0	0	0	0	0	0	2	-
Total	88.2	11	4	6	4	3	4	2	

Table 10 : Results of Stepwise Discriminant Analysis for Data Set 3

(A) CLASSIFICATION FUNCTIONS

Variable	Group							
	CO	DA	RA	PT	PM	CR	BK	TK
E2U	3.125	3.210	3.401	2.901	3.237	2.899	3.209	3.587
STU	10.396	11.287	11.530	11.531	10.884	11.096	11.307	13.067
Constant	-411.043	-472.075	-500.601	-470.924	-448.105	-441.653	-473.304	-622.841

(B) CLASSIFICATION MATRIX

Group	% correctly classified	No. of cases classified into group								Within group Sl. Nos. of cases misclassified
		CO	DA	RA	PT	PM	CR	BK	TK	
CO	90.0	9	0	0	0	1	0	0	0	1
DA	0.0	1	0	2	0	1	0	1	0	BK-1,RA-2,CO-3,RA-4,PM-5
RA	60.0	0	0	3	0	2	0	0	0	1,3
PT	60.0	0	0	0	3	0	1	1	0	BK-4,CR-5
PM	50.0	1	0	1	0	2	0	0	0	CO-2,RA-4
CR	0.0	0	0	0	3	2	0	0	0	PT-1,PT-2,PT-3 PM-4,PM-5
BK	66.7	0	0	0	0	1	0	2	0	1
TK	100.0	0	0	0	0	0	0	0	2	-
Total	53.8	11	0	6	6	9	1	4	2	

was also calculated for each individual in Data sets 2 and 3, but because of the limited use to which these probabilities can be put to, we have not presented these values.

10. Discussion

From the foregoing analysis we find several significant features; these are discussed below:

The results of the cluster analyses show that, by and large, the quantitative flexion/extension and stance variables are useful in characterising the various groups, and that the grouping of individuals based on these quantitative variables agree with the grouping based on clinico-pathological observations. The only exception seems to be when all the 8 groups were clustered on the basis of measurements on 7 common variables (figure 3). This, however, is not very unexpected since all the 7 variables pertained to the unaffected leg (because of the very nature of the orthopaedic impairment in some of the groups, e.g., the crutch users) and it is natural that we could not distinguish the various orthopaedically handicapped groups from the control group.

The results of the analyses of variance and discriminant analysis show that the stance and knee variables are more important for characterising the groups than the ankle variables. There are several reasons for this pattern. First, stance is one of the primary variables for characterisation of gait; the flexion/extension variables are only secondary. Of the stance variables, STU seems to be a better discriminator than STA. This is also expected. Orthopaedic handicap does affect STA, but it affects STU to a greater extent. This is simply because of the biological fact that one wants to reduce the time during which the entire weight of the body is carried on the affected leg during walking, and this leads to an increase in the length of time when the unaffected foot is on the ground in comparison with a control individual, thereby causing an increase in the value of STU. Second, the reason why in this particular study the knee variables turned out to be of greater

importance than ankle variables is, perhaps, simply because of the fact that most of the orthopaedically impaired individuals considered in this study have knee, and not ankle, problems. The fact that of the ankle variables, F3 turned out to be the only variable of importance is in conformity with medical expectations.

That only 3 or 4 of the considered variables yields a misclassification probability of only about 0.1, except for Data Set 3, is extremely encouraging. In the CR group one lower limb is either amputated or paralysed, the other limb having no orthopaedic impairment. From the clinical viewpoint the limb may be classed as "unaffected". During walking with crutches however, the kinematics of this limb is quite different from an unaffected normal limb. The body swings forward from the support at the axilla, with much less flexions at the knee and ankle. For the purpose of gait mechanics the limb is thus an "affected" limb. The inclusion of this biomechanically affected leg into the unaffected group has probably led to the 100% misclassification of the CR group in data set 3. An examination of the posterior probabilities suggest that the 'Post-Menisectomy' individuals are rather close to the Control individuals, and that there is a large degree of overlap in the characteristics of the two groups of individuals with arthritis - OA and RA. Although these are not presented, a study of the posterior probabilities (computed on the basis of observations on 10 variables) indicates that the amputees - BK and TK, are quite dissimilar in comparison with control individuals. These features, by and large, agree with medical observations. The ranking of individuals in the order of their similarity to the controls based on medical case-histories do not, however, correspond too well with the ranking obtained by the use of posterior probabilities.

11. Suggestions

1. In order to obtain estimates of intra-individual variation, observations on 20 strides should be kept separately, and should not be averaged (as has been done with the data analysed in this study).

2. Sample sizes for each group should be increased (to at least about 50 individuals per group) in order that the effects of variables like age, sex, etc., can be studied. Obviously, the greater the variation in covariates, e.g., age, body weight, speed of walking, etc., the greater should be the sample size .

12. Statistical Treatment of Some Miscellaneous Questions

Among the Control individuals it is seen from the data that there is very little difference in the measurements on a variable for the two legs. There is, however, a great deal of variation between legs for individuals belonging to the orthopaedically handicapped groups. The question then arises: Can we discriminate between the groups better if we consider the differences between the observations of the two legs for each variable (e.g., E1) instead of treating the observations on this variable for the two legs separately (e.g., E1U and E1A) ? To check this, we defined variables as : E1 = E1U - E1A, E2 = E2U - E2A, F1 = F1U - F1A, ... , F4 = F4U - F4A, and ST = STU - STA. With these 7 new variables, we performed a discriminant analysis taking only the 5 groups (C0, DA, RA, PT and PM), for obvious reasons. The results obtained are as follows : (i) There are three variables that are significant for purposes of discrimination, which are, in the order of importance, E2, ST and F3. (ii) The percent of correct classification using the estimated classification functions is 86.2%. These results show that there is no betterment in discriminatory ability by considering the newly defined variables. However, it is interesting to note that in the corresponding analysis with original variables none of the ankle variables entered the discriminant functions. In the present case, of the three variables entered, one is a knee variable, the second is a stance variable and the third is an ankle variable.

ıtra-individual variation,
ärately, and should not
analysed in this study).
ɒe increased (to at least
the effects of variables
the greater the variation
valking, etc., the greater

Questions

en from the data that
nents on a variable for
al of variation between
ally handicapped groups.
etween the groups better
bservations of the two
reating the observations
g., E1U and E1A) ? To
E1A, E2 = E2U - E2A,
= STU - STA. With these
analysis taking only the
us reasons. The results
ariables that are signifi-
the order of importance,
sification using the esti-
results show that there
considering the newly
note that in the corres-
the ankle variables en-
case, of the three vari-
nd is a stance variable

Since one of the objectives of this study was to find an index of similarity of an individual to a control individual, we have proposed that the posterior classification probability be used as an index. However, in computing and illustrating the use of the posterior probability as an index of similarity (see, page 37 and table 6), we have assumed that the individual can come from either the CO group or from one of several, say k, orthopaedically handicapped groups. The prior probabilities were assumed to be 1/(k+1) for an individual's inclusion in one of the (k+1) groups. However, realistically it may be more useful to only consider two groups CO and $\overline{CO}$, where CO is the group of all individuals not belonging to the control group. In other words, it may be more realistic to just say that an individual is either orthopaedically impaired or has no orthopaedic impairment.

We, therefore, performed discriminant analyses and calculated the posterior probabilities by reducing the data as :

(i) Reduced Data Set 1 :
Groups : CO and $\overline{CO}$ = (DA, RA, PT, PM)
Variables = 14
Sample sizes : CO = 10, $\overline{CO}$ = 19

(ii) Reduced Data Set 2 :
Groups : CO and $\overline{CO}$ = (DA, RA, PT, PM, BK, TK)
Variables = 10
Sample sizes : CO = 10 and $\overline{CO}$ = 24

The results for the Reduced Data Set 1 are presented in Table 11. As is seen from this table, the variables entering the classification functions are STU and E1A. If we compare these results with those for Data Set 1 given in tables 5 and 6, we find that there is only one common discriminating variable - STU. The probability of correct classification is the same for both these data sets - 0.897. The individuals misclassified are, however, different for these two data sets. The posterior

Table 11 : Results of Stepwise Discriminant Analysis for Reduced Data Set 1

(A) CLASSIFICATION FUNCTIONS

Variable	Group	
	CO	CO
STU	6.907	7.686
EIA	0.732	0.311
Constant	-212.496	-255.375

(B) POSTERIOR PROBABILITIES

Group	Sl. No.	Posterior Probability for Group		Misclassified to Group
		CO	CO	
CO	1	0.988	0.012	
	2	0.964	0.036	
	3	0.991	0.009	
	4	0.986	0.014	
	5	0.990	0.010	
	6	0.998	0.002	
	7	0.934	0.066	
	8	0.924	0.076	
	9	0.996	0.004	
	10	0.920	0.080	
DA	1	0.000	1.000	
	2	0.000	1.000	
	3	0.557	0.443	CO
	4	0.003	0.997	
	5	0.021	0.979	
RA	1	0.133	0.867	
	2	0.001	0.999	
	3	0.104	0.896	
	4	0.001	0.999	
	5	0.001	0.999	
PT	1	0.137	0.863	
	2	0.002	0.998	
	3	0.007	0.993	
	4	0.019	0.981	
	5	0.000	1.000	
PM	1	0.388	0.612	
	2	0.920	0.080	CO
	3	0.733	0.267	CO
	4	0.025	0.975	

Analysis for Reduced

oup
--- Misclassified
to Group

CO

CO
CO

probabilities are also different, and the posterior probability of an individual being classified into the CO group has increased for the Reduced Data Set 1 in comparison with Data Set 1. This is, of course, expected since in the Reduced Data Set we have pooled all the 4 distinct orthopaedically impaired groups.

The results for Reduced Data Set 2 are qualitatively similar. The variables entering the discriminant functions are, in order of importance, STU, F4U, STA and E1U. The probability of correct classification is 0.882.

13. References

Bhadra, N. and Bose, K.S. (1983) : Sagittal plane knee movement during level walking. Ind. J. Surg. (in Press.)

Lamoreux, L.W. (1971) : Kinematic measurements in the study of human walking. Bull. Prosth. Res. 10 : 3-84.

Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979) : Multivariate Analysis. New York : Academic Press.

Rao, C.R. (1952) : Advanced Statistical Methods for Biometric Research. New York : John Wiley & Sons.

Saunders, J.B. de C., Inman V.T. and Eberhart, H.D. (1953) : Major determinants in normal and pathological gait. J. Bone and Joint Surg. 35-A : 543-558.