

A comparison of within- and between-day reliability of discrete 3D lower extremity variables in runners

R. Ferber ^{a,*}, I. McClay Davis ^{a,b}, D.S. Williams III ^c, C. Laughton ^d

^a Motion Analysis Laboratory, Department of Physical Therapy, University of Delaware, 301 McKinly Lab, Newark, DE 19716-2591, USA

^b Joyner Sportsmedicine Institute, Harrisburg, PA 17111, USA

^c Department of Physical Therapy, East Carolina University, Greenville, NC 27858, USA

^d Department of Biomedical Engineering, Boston University, Boston, MA, USA

Received 24 October 2001; accepted 12 April 2002

Abstract

It is important to understand the day-to-day variability that is attributed to repositioning of markers especially when assessing a treatment effect or response over time. While previous studies have reported reliability of waveform patterns, none have assessed the repeatability of discrete points such as peak angles, velocities and angular excursions which are often used when making statistical and clinical comparisons. The purpose of this study was to compare the within- and between-day variability of discrete kinematic, kinetic, and ground reaction force (GRF) data collected during running. Comparisons for 20 recreational runners were evaluated for within- and between-day reliability of discrete 3D kinematic, kinetic, and GRF variables. The results indicated that within-day comparisons were more reliable than between-day. Joint angular velocity and angular excursion values were more reliable between-days as compared to absolute peak angle measures and may be more useful in interpreting changes in treatment over time. Between-day kinematic and kinetic sagittal plane values were more reliable than secondary plane values. Reliability of GRF data was greater than kinematic and kinetic data for between-day comparisons.

© 2002 Orthopaedic Research Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Gait analysis; Repeatability; Kinematics; Kinetics

Introduction

Three-dimensional motion analysis has become a widely used tool for both research and clinical assessments. However, in order for motion analysis to be valuable, one needs to have reliable data. One commonly recognized problem is the day-to-day variability that may be present due to placement of markers over the skin. This variability is especially important when the same subject is being tested on more than one occasion and comparisons are made between the sessions particularly when determining a clinical or surgical outcome on a pre/post-treatment basis.

Relatively few investigations have been conducted to assess reliability of kinematics, kinetic, and ground reaction force (GRF) variables during human movement [1,3,5–9]. One of the first reliability investigations was by

Kadaba et al. [3] who used coefficients of multiple correlations to compare within- and between-day repeatability of kinematic and kinetic waveforms during human gait. Four important results were reported. First, repeatability for all parameters was better within a test day (between trials) as compared to consecutive testing sessions. Carson et al. [1] supported these findings and reported between-day variability was greater than within-day variability for foot kinematic patterns for two subjects during normal gait.

Second, Kadaba et al. [3] identified a constant offset in joint angle between testing days. When data were adjusted to account for this offset, between-day repeatability improved significantly but was still less reliable than within-day values. Carson et al. [1] also reported an overall shift in the absolute value of a joint angle between testing sessions and suggested that the shift could be reduced by subtracting out a neutral position obtained during the static trial. A shift in the data between-days would likely have an effect on absolute measures such as peak values to a greater degree than relative

* Corresponding author. Tel.: +1-302-831-4646; fax: +1-302-831-4234.

E-mail address: reedferb@udel.edu (R. Ferber).

measures such as excursions suggesting the latter may be more reliable between-days.

Third, Kadaba et al. [3] demonstrated relatively low reliability in frontal and transverse plane values as compared to sagittal plane values in both within- and between-day comparisons. The greater variability was in the transverse plane which the authors attributed to the alignment of the wand markers. In a study comparing various marker configurations to bone anchored markers on the tibia, Manal et al. [6] reported that the greatest error between bone and skin markers was in the transverse plane. These data suggest that soft tissue movement may be greatest in the secondary plane as compared to the sagittal plane and may significantly influence both within- and between-day variability.

Finally, Kadaba et al. [3] reported that GRF data demonstrated greater repeatability as compared to kinematic values but offered no explanation for this result. Winter [9], finding similar results, suggested that since GRF data was representative of the sum of all the segmental masses and accelerations, and less variability would be seen as compared to individual joint kinetic or kinematic patterns.

While these investigations provide important information regarding the repeatability of measurement of movement patterns, discrete points such as peak angles, velocities and angular excursions are often the variables of interest used when making statistical and clinical comparisons. It is important to understand the magnitude of change that is attributed to repositioning of markers especially when assessing a treatment effect or response over time. As well, previous studies have focused on the reliability of gait variables during walking. Reliability of these variables may be different during running due to the greater forces and joint excursions involved. Therefore, the purpose of this study was to compare the within- and between-day variability of discrete kinematic, kinetic, and GRF data collected during running.

Four hypotheses were formulated based on the review of literature: (1) between-day variability would be greater as compared to within-day variability, (2) kinematic angular excursion and angular velocity values would be less variable than peak angle values, (3) sagittal plane kinematics and kinetics would be less variable than frontal and transverse plane values and, (4) GRF data would be less variable than joint kinetic and kinematic data.

Materials and methods

Twenty uninjured recreational runners, 7 males and 13 females, volunteered for this study. The mean age, body mass, and body height of subjects were 21.4 yr (± 3.1 yr), 67.0 kg (± 11.7 kg), and 173.1 cm (± 9.3 cm), respectively. Prior to participation, each subject signed a consent form approved by the University's Human Subjects Compliance Committee.

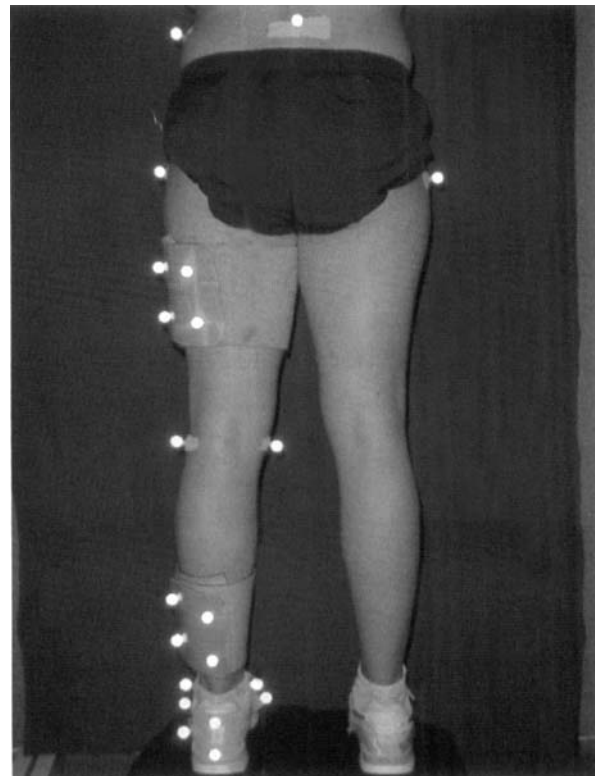


Fig. 1. Retroreflective marker placement on the tested lower extremity.

Retroreflective markers were positioned on specific anatomical landmarks of the pelvis, thigh, shank and foot (Fig. 1). Subjects ran along a 25 m runway at a speed of 3.65 m/s ($\pm 5\%$). Running speed was monitored using photoelectric cells placed 2.86 m apart along the runway. Running trials were collected for the right lower extremity during stance. Subjects returned one week later in which the same procedure was employed. The same tester attached the markers on all subjects.

Kinematic data were collected with a passive, 6-camera, 3D motion analysis system (VICON, Oxford Metrics, UK). Kinematic data were sampled at 120 Hz and low-pass filtered at 8 Hz with a fourth-order zero lag Butterworth filter. Kinetic GRF data were collected using a force plate (BERTEC Corp., Worthington, OH, USA). GRF data were collected at 960 Hz and low-pass filtered at 50 Hz with a fourth-order zero lag Butterworth filter. Trials were normalized to 100% of stance and five were averaged for each subject. Force data were normalized to body weight, and moment and power data were normalized to body mass and height.

MOVE3D software (MOVE3D, NIH Biomotion Laboratory, Bethesda, MD, USA) was used to calculate kinematic and kinetic variables. All lower extremity segments were modeled as a frustra of right cones model and anthropometric data provided by Dempster [2]. Variables of interest included lower extremity joint (ankle, knee, and hip) peak angle, angular excursion, angular velocity, internal moments, power, 3D GRF data, and running speed. Variables were selected from heel strike to the first 60% of stance.

Intraclass correlation coefficients (ICC) were used for the variables of interest to compare within- (ICC (2,1)) and between-day (ICC (2,k)) reliability. To test the hypotheses, 95% confidence intervals (CI) were calculated from the ICC values and evaluated for overlap within comparisons.

Results

The mean and standard error of measurement (SEM) for the kinematic, kinetic, and GRF variables of interest

Table 1
Within- and between-day mean, SEM, and ICC kinematic joint peak angle, angular excursion, and angular velocity values

		Day 1		Day 2		Mean difference
		Mean (SEM)	ICC within-day	Mean (SEM)	ICC between-day	
<i>Peak angle (°)</i>						
Ankle	Dorsiflexion	24.68 (0.81)	0.92	24.24 (1.03)	0.85	0.44
	Eversion	7.70 (0.49)	0.96	7.88 (2.10)	0.63	0.18
	Abduction	-7.57 (0.32)	0.98	-6.68 (2.11)	0.68	0.89
Knee	Flexion	-44.99 (1.00)	0.92	-45.64 (2.22)	0.93	0.65
	Adduction	1.16 (0.04)	0.98	0.52 (2.31)	0.71	0.64
	Int. rot.	2.70 (0.03)	0.98	1.14 (1.63)	0.83	1.56
Hip	Extension	36.69 (0.98)	0.98	37.67 (2.21)	0.88	0.98
	Adduction	8.43 (0.32)	0.99	10.07 (2.70)	0.69	1.64
	Int. rot.	10.50 (0.35)	0.98	10.35 (3.20)	0.54	0.15
<i>Peak excursion (°)</i>						
Ankle	Dorsiflexion	19.61 (1.05)	0.88	18.92 (1.48)	0.83	0.69
	Eversion	14.67 (1.28)	0.89	15.05 (1.10)	0.93	0.38
	Abduction	-8.26 (0.78)	0.92	-8.74 (0.97)	0.91	0.48
Knee	Flexion	-40.94 (1.67)	0.90	-40.05 (1.16)	0.95	0.89
	Adduction	3.97 (0.14)	0.96	3.64 (1.73)	0.68	0.33
	Int. rot.	10.06 (0.89)	0.97	10.09 (1.13)	0.94	0.02
Hip	Extension	-0.41 (0.05)	0.82	-0.38 (0.03)	0.88	0.03
	Adduction	6.59 (0.85)	0.88	6.03 (1.11)	0.84	0.56
	Int. rot.	-0.01 (0.01)	0.87	-0.01 (0.01)	0.82	0.00
<i>Peak velocity (°/s)</i>						
Ankle	Dorsiflexion	291.69 (12.18)	0.90	284.79 (9.25)	0.94	6.89
	Eversion	204.49 (21.32)	0.90	215.18 (22.65)	0.90	10.69
	Abduction	-153.48 (15.02)	0.91	-155.29 (18.95)	0.86	1.81
Knee	Flexion	-527.57 (18.48)	0.92	-529.62 (12.38)	0.96	2.05
	Adduction	158.02 (10.11)	0.97	161.71 (26.78)	0.84	3.69
	Int. rot.	104.74 (11.43)	0.94	103.28 (12.35)	0.94	1.47
Hip	Extension	102.65 (8.86)	0.89	110.32 (31.88)	0.64	7.67
	Adduction	109.52 (8.40)	0.94	105.05 (11.25)	0.86	4.47
	Int. rot.	107.32 (15.72)	0.91	95.11 (17.16)	0.88	12.21

along with within- and between-day ICC values are shown in Tables 1 and 2. The comparisons of 95% CI for each hypothesis can be seen in Fig. 2 revealing that three of the four hypotheses were supported. As expected, between-day ICC values were less reliable than within-day ICC values (Fig. 2). Observation of Tables 1 and 2 indicated that between-day ICC values < 0.70 were more frequent than within-day variables. For those variables that exhibited between-day ICC values < 0.70, corresponding within-day ICC values were > 0.87.

Surprisingly, relative peak angular velocity and excursion values were not less variable than absolute peak angle values when averaging across within- and between-day conditions (Fig. 2). However, when assessing the between-day condition alone, peak angular excursion and velocity ICC values were higher than their respective peak angle ICC values as initially hypothesized

(Table 1). The greatest individual between-day mean differences for peak angle and the corresponding peak angular excursion values were determined and are presented in Table 3. The results indicated that the average difference in peak angle was 9.04° whereas the average difference in the corresponding angular excursion was 3.63°.

In support of the hypothesis, sagittal plane ICC values were less variable than frontal and transverse plane ICC values, the latter two being similar to each other (Fig. 2). Observation of Tables 1 and 2 reveal that the strength of this finding was in the between-day results as the ICC values were remarkably similar across planes in the within-day data.

As expected, GRF variables were more reliable than the kinematic and kinetic values (Fig. 2). Again, these findings were more attributed to the between-day

Table 2
Within- and between-day mean, SEM, and ICC joint peak moment and power values, GRF peak values, and running speed

		Day 1		Day 2		Mean difference
		Mean (SEM)	ICC within-day	Mean (SEM)	ICC between-day	
<i>Peak moment (Nm/(kg ht))</i>						
Ankle	Plantarflexion	-1.48 (0.29)	0.95	-1.47 (0.40)	0.94	0.01
	Inversion	-0.12 (0.01)	0.97	-0.12 (0.11)	0.73	0.01
	Adduction	0.01 (0.02)	0.85	0.01 (0.06)	0.42	0.00
Knee	Extension	1.23 (0.27)	0.86	1.22 (0.46)	0.80	0.02
	Abduction	-0.50 (0.04)	0.98	-0.44 (0.10)	0.98	0.06
	Ext. rot.	-0.08 (0.01)	0.97	-0.05 (0.06)	0.73	0.03
Hip	Extension	-1.36 (0.10)	0.92	-1.28 (0.34)	0.95	0.07
	Abduction	-1.28 (0.18)	0.96	-1.28 (0.45)	0.91	0.01
	Ext. rot.	0.32 (0.12)	0.73	0.34 (0.13)	0.73	0.02
<i>Peak power (W/(kg ht))</i>						
Ankle	Plantarflexion	-5.07 (1.47)	0.86	-4.90 (1.68)	0.92	0.18
	Inversion	-0.01 (0.01)	0.85	-0.02 (0.04)	0.75	0.01
	Adduction	-0.10 (0.01)	0.87	-0.11 (0.11)	0.75	0.01
Knee	Extension	2.39 (0.36)	0.95	2.74 (0.52)	0.93	0.35
	Abduction	0.17 (0.02)	0.94	0.17 (0.02)	0.87	0.01
	Ext. rot.	-0.11 (0.01)	0.88	-0.07 (0.06)	0.82	0.04
Hip	Extension	3.56 (0.74)	0.85	3.93 (0.98)	0.81	0.37
	Abduction	-1.45 (0.07)	0.91	-1.12 (1.02)	0.67	0.33
	Ext. rot.	-0.32 (0.01)	0.86	-0.34 (0.23)	0.87	0.02
<i>GRF (BW)</i>						
	Lateral	0.09 (0.01)	0.94	0.09 (0.01)	0.96	0.00
	Medial	-0.09 (0.01)	0.92	-0.05 (0.04)	0.91	0.02
	Braking	-0.37 (0.02)	0.88	-0.44 (0.07)	0.91	0.03
	Propulsion	0.35 (0.02)	0.88	0.39 (0.04)	0.93	0.03
	Vertical	2.60 (0.07)	0.90	2.56 (0.04)	0.95	0.05
Running speed (m/s)		3.57 (0.01)	0.99	3.62 (0.01)	0.98	0.05

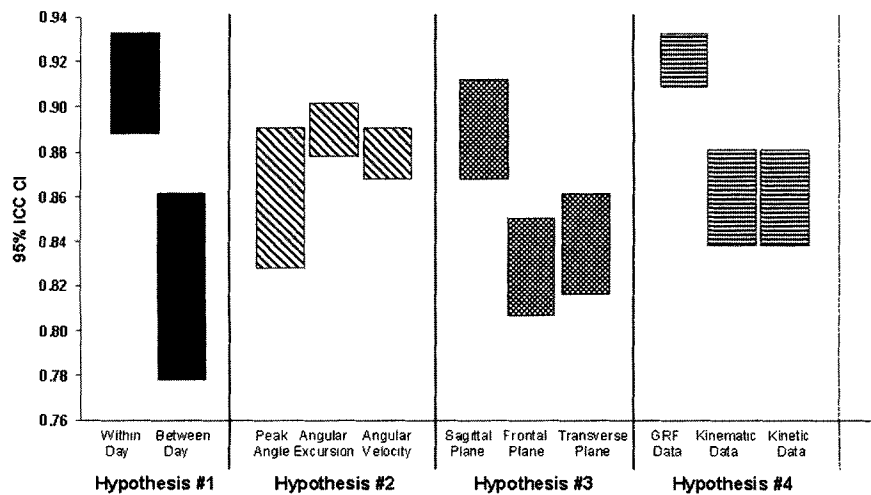


Fig. 2. 95% CI of selected variables to test the four hypotheses.

condition as the GRF ICC values were surprisingly similar to kinematic and kinetic values for the within-

day condition (Table 2). Running speed was reliable both within and between testing sessions (Table 2).

Table 3

Greatest individual subject between-day mean difference of peak angle and the associated angular excursion value

		Subject	Peak angle (°)			Peak excursion (°)		
			Day 1	Day 2	Difference	Day 1	Day 2	Difference
Ankle	Dorsiflexion	1	22.90	15.57	7.33	21.10	21.90	0.80
	Eversion	14	1.20	8.92	7.72	14.75	22.10	7.35
	Abduction	20	-12.80	-4.03	8.77	-7.52	-4.63	2.89
Knee	Flexion	16	-52.80	-38.80	14.00	-51.90	-49.30	2.60
	Adduction	19	4.39	10.50	6.11	4.52	1.22	3.30
	Int. rot.	4	0.89	6.93	6.04	9.55	12.50	2.95
Hip	Extension	8	26.20	39.10	12.90	-0.42	-1.80	1.38
	Adduction	11	11.40	28.30	16.90	14.00	4.02	9.98
	Int. rot.	13	5.54	13.10	7.56	6.55	5.14	1.41

Discussion

The purpose of this study was to compare the within- and between-day variability of discrete kinematic, kinetic, and GRF data collected during running. The results of this investigation demonstrate that most variables of interest exhibited high repeatability both within and between testing sessions and overall, three of the four stated hypotheses were supported.

It has been reported that alterations in walking and running speed can result in significant changes in peak joint moment and power values, particularly at the knee and hip joint [4,9]. In the present investigation, running speed was controlled over a narrow range (3.46–3.83 m/s) but variations in running speed could result in some within- and between-day variability of the selected variables. However, since running speed was one of the most consistent variables analyzed, it is unlikely that other variables of interest were significantly influenced by variations in running speed between trials or between testing sessions (Table 2).

In the first comparison, within-day kinematic and kinetic ICC values were higher as compared to between-day values which are in agreement previous investigations [1,3,8]. Within-day variability has been attributed to measurement error, skin marker movement, and inherent physiological variability during human locomotion [1,3,5,8]. In the present investigation, calibration errors were held below 3 mm. As well, all cameras were linearized and camera positions and parameters were optimized to allow for the least possible measurement error. However, since it is difficult to separate error due to the measurement system and error due to external factors, the variability reported in this study include their contribution. Artifact due to soft tissue movement has been reported to produce secondary plane movement errors of 60–70% relative to the range of motion during the stance phase of running [6,7]. However, the good within-day repeatability observed in the present

investigation suggests that inaccuracies due to skin movement and movement variability were repeatable and systematic.

Between-day kinematic and kinetic ICC values were lower as compared to within-day values as several variables fell below an ICC value of 0.75. In particular, 6 of 9 peak angle, and 7 of 18 joint moment and power variables exhibited significant declines from within- to between-day repeatability. These results suggest that between-day variability can be attributed to factors affecting within-day variability as well as marker reapplication error [1,3,8]. Anatomical marker placement is especially important to reliability as this establishes the anatomical coordinate system (ACS) about which axes the angles are decomposed. Slight changes in these marker positions may result in producing cross-talk between the planes of motion, or simply producing an offset shift in the data, resulting in lower between-day variability. To reduce this variability, a single, well-trained investigator applied all markers on successive sessions. However, lower between-day ICC values suggest that differences in marker placement influenced measurement repeatability regardless of controlling for intertester variability.

While between-day ICC values were somewhat lower as compared to within-day values, little difference was seen when comparing the actual data from day-to-day (Tables 1 and 2). For example, the between-day ICC for rearfoot eversion was only 0.63 while the mean values for these variables were very similar (7.70° vs. 7.88° respectively) and the SEM values were relatively higher between-days (0.49° vs. 2.10° respectively). These low ICC and higher SEM scores reflect individual subject differences between testing sessions. However, it appears that these differences were equally and randomly distributed across the subjects resulting in similar mean data. It has also been suggested that variables having small means will be associated with greater variability [8]. Nonetheless, these results suggest that one may have

confidence when comparing mean group data between testing sessions but should exercise care when making individual comparisons.

Contrary to the hypothesis, relative peak angular velocity and excursion values were not less variable than absolute peak angle values (Table 1). However, upon closer inspection, the hypothesis was supported for the between-day condition and it was the higher than expected within-day reliability of the peak angles that caused the hypothesis to be rejected. Observation of Table 1 demonstrates that peak angle, angular velocity, and angular excursion variables all exhibited within-day ICC values > 0.82 and relatively low SEM values. Since within-day repeatability results are not affected by errors in marker reapplication, the high within-day ICC and low SEM values indicate that the subjects ran in a very similar manner across trials. However, between-day ICC values for 6 of 9 peak angle values were below 0.72 as compared to only 1 excursion and 1 velocity variable (Table 1). These results are in agreement with those of Kadaba et al. [3] and Carson et al. [1] who reported that misalignment of markers introduced a constant offset in joint angle between testing days and lower repeatability of absolute kinematic measures. Since peak values are absolute measures, they are more likely affected by alterations in the ACS alignments than the relative measures of velocity and excursion. In all but one case there was a substantial reduction in the day-to-day differences when using excursions rather than peak values when examining individual differences between testing days (Table 3). These results suggest that relative measures such as joint angular excursion and velocity are more reliable than absolute measures such as joint angle when assessing a subject's running mechanics over time.

Overall, ICC values for secondary plane variables were lower than sagittal plane variables (Tables 1 and 2). However, as with the previous hypothesis, this was most notable in the between-day condition. The high ICCs for the secondary plane variables in the within-day data were surprising and in contrast with previous literature [1,3,5,8]. These investigations suggested that variability in walking patterns was mostly reflected in the transverse plane motions. It is possible that the dynamic and ballistic nature of running results in greater within-day repeatability in the secondary planes of motion than walking. As well, the tibial marker set used in the present investigation was placed on the distal-lateral position of the tibia (Fig. 1). This marker position has been reported to yield the best estimate of tibial rotation as compared with bone-mounted markers [6]. This may have also attributed to the relatively high within-day secondary plane ICC values.

The between-day sagittal plane kinematic and kinetic values were generally less variable as compared to secondary plane variables (Tables 1 and 2). These results

are similar to previous investigations and suggest that variability due to marker reapplication is less evident in the sagittal plane [1,3,5,8]. Care in marker placement must be taken to minimize between-day variability of secondary plane kinematic and kinetic variables.

As expected, GRF values were more reliable than joint kinematic or kinetic data (Table 2). Winter [9] suggested that GRF values, being the sum of all the segmental masses and accelerations and gravitational forces, would be less variable than joint kinematic or kinetic data. In addition, no markers are necessary to collect GRF data and would therefore be less variable. Within-day ICCs were surprisingly similar between kinematic, kinetic, and GRF variables. However, most kinematic and kinetic variables exhibited high repeatability across trials suggesting the subjects ran consistently and repeatably across trials as indicated by whole-body or individual joint movement patterns. In partial support of the hypothesis, between-day GRF variables exhibited greater repeatability than kinematic and kinetic values (Table 2). These results are similar to those of Kadaba et al. [3] and Winter [9] and suggests that GRF variables exhibit high within- and between-day repeatability.

As previously noted, variability in kinematic and kinetic variables may be attributed to measurement error, skin marker movement, marker reapplication errors, and inherent physiological variability during human locomotion. While the first three factors are independent of the patient population itself, it is possible that physiological variability may be different in a clinical population. Steinwender et al. reported that children with cerebral palsy exhibited lower within-day repeatability of kinetic and kinematic variables but greater between-day kinetic repeatability [8]. However, changes in variability in adult clinical populations, as compared to pretreatment status or compared to a healthy population, are unknown. Future studies addressing within- and between-day reliability of involving clinical patients using motion analysis populations would help to answer this question.

In summary within-day comparisons were more reliable than between-day comparisons. Within-day variability was similar between planes of motion while between-day sagittal plane ICC values were less variable as compared to secondary plane variables. Angular velocity and excursion ICC values were more reliable from day to day and these values may be more useful than peak joint angles in interpreting changes in treatment over time. GRF data were more reliable between testing sessions as compared to kinematic and kinetic data and suggests that these variables may be most reliable when assessing between-day differences. These types of studies should be performed by investigators before undertaking studies which involve data collection on the same subject over time.

Acknowledgements

We would like to thank Tracy A. Dierks and Robert J. Butler for their assistance in data processing.

References

- [1] Carson MC, Harrington ME, Thompson N, O'Connor JJ, Theologis TN. Kinematic analysis of a multi-segment foot model for research and clinical applications: a repeatability analysis. *J Biomech* 2001;34(10):1299–307.
- [2] Dempster W. Space requirements of the seated operator. In: WADC Technical Report Ohio:L Wright–Patterson Air Force Base, 1959. p. 55–159.
- [3] Kadaba MP, Ramakrishnan HK, Wootten ME, Gainey J, Gorton G, Cochran GV. Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. *J Orthop Res* 1989;7(6): 849–60.
- [4] Kyrolainen H, Belli A, Komi PV. Biomechanical factors affecting running economy. *Med Sci Sports Exerc* 2001;33(8):1330–7.
- [5] Liu W, Siegler S, Hillstrom H, Whitney K. Three-dimensional, six-degrees-of-freedom kinematics of the human hindfoot during the stance phase of level walking. *Human Movement Sci* 1997;16: 283–98.
- [6] Manal K, McClay I, Stanhope S, Richards J, Galinat B. Comparison of surface mounted markers and attachment methods in estimating tibial rotations during walking: an in vivo study. *Gait Posture* 2000;11(1):38–45.
- [7] Reinschmidt C, van den Bogert AJ, Nigg BM, Lundberg A, Murphy N. Effect of skin movement on the analysis of skeletal knee joint motion during running. *J Biomech* 1997;30(7): 729–32.
- [8] Steinwender G, Saraph V, Scheiber S, Zwick EB, Uitz C, Hackl K. Intrasubject repeatability of gait analysis data in normal and spastic children. *Clin Biomech* 2000;15(2):134–9.
- [9] Winter DA. Kinematic and kinetic patterns in human gait: variability and compensating effects. *Human Movement Sci* 1984; 3:51–76.