IDENTIFICATION OF MOVEMENT STRAGEGIES IN VERTICAL JUMPS

Chris Richter[1, 2, 3], Brendan Marshall[1, 2], Noel E. O'Connor[1] and Kieran Moran[1, 3]

Insight Centre for Data Analytics, Dublin City University, Ireland[1]
Department of Sport Medicine, Sports Surgery Clinic, Dublin, Ireland[2]
School of Health and Human Performance, Dublin City University, Dublin, Ireland[3]

The primary aim of this study is to compare the ability of three commonly used clustering techniques to identify movement strategies within countermovement jumps. A secondary aim is to interpret the identified movement strategies. A hierarchical, k-means using non- and normalized subject scores and an Expectation-Maximization approach using normalized subject scores were examined. The ability to identify movement strategies was measured using the r^2-value of a regression model to describe jump height. Clusters of the best clustering solution were examined for differences. Hierarchical clustering utilizing normalized subject scores to generate 4 clusters appears to be the most suitable technique. The generated clusters demonstrated clear defining characteristics.

KEY WORDS: single group analysis, subgroup analysis, continuous data analysis, performance related factors, countermovement jump

INTRODUCTION: A major aim of biomechanical studies is the identification of performance related features within a movement to allow development of high quality training programs. To identify such factors, biomechanical studies commonly use a single group analysis design assuming that performance related features are equal across individuals. However, different individuals may use different movement strategies (Vanezis & Lees, 2005), and in turn may have different performance related features. When examining a group of individuals that use different movement strategies, a single group design can mask performance related factors (Stergiou & Scott, 2005; Nicholas Stergiou, 2004). A solution to a single group analysis design is the subgroup analysis design. A subgroup analysis design clusters individuals with similar movement patterns into separate groups (and allows the identification of performance related factors for each of the movement strategies with the examined group). The advantage of the subgroup analysis design have been demonstrated in human gait studies (e.g. Toro, Nester, & Farren, 2007). However, when performing a subgroup analysis design the user has to select a clustering technique and the number of movement strategies that have to be separated. While a number of clustering techniques exist, which may result in different clusters (Jain, Murty, & Flynn, 1999; Martinez, Martinez, & Solka, 2004), there is a lack of biomechanical studies that have compared the ability of cluster techniques to identify movement strategies. Hence, the primary aim of this study is to compare the ability of three commonly used clustering techniques to identify movement strategies within the countermovement jump (CMJ). A secondary aim is to interpret the identified movement strategies.

METHODS: This study recruited 122 athletes, who were free from any injury and experienced in performing a CMJ. The University Ethics Committee approved the study and all subjects signed an informed consent form before participation. Prior to data collection, every subject completed a standard warm-up routine. The subjects performed 15 maximum effort CMJs without an arm swing, standing with each foot on a force platform, and rested for 30 seconds between the trials. A motion analysis system (Vicon 512 M, Oxford Metrics Ltd, England) and two force plates (BP-600900, AMTI, MA, USA) recorded the position of a set of spherical reflective markers (250 Hz) and the vertical ground reaction force (1000 Hz), respectfully. Jump height was calculated by the center of mass velocity at takeoff. Based on jump height, the best jump performance of each subject was chosen for data analysis. All curves were normalized to body mass and only the propulsion phase was used for analysis because the performance outcome (jump height) is fully determined by the propulsion phase (impulse-momentum relationship). Kinematic and kinetic variables for each joint were

computed for the left and right body side. The average of a kinematic and kinetic variable from both body sides was used for data analysis. Joint kinetics were calculated using inverse dynamics (Winter, 2009). To cluster the captured kinematic and kinetic waveforms, subject scores were computed using Analysis of Characterising Phases (Richter et al., 2013a). Subject scores were computed for key phases using the magnitude domain. Key phases were identified using VARIMAX rotated functional principal components, which retained more than 99 % of the variance within the data's magnitude domain (Richter et al., 2013b). To classify the data, the computed subject scores were input into a hierarchical and a k-means approach using normalized and non-normalized subject scores, and into an Expectation-Maximization algorithm using non-normalized subject scores. The normalization of the input data was performed by transforming the subject scores into their correlation matrix, which quantifies numerically the relationship between them. The hierarchical algorithm calculated pairwise distances using Euclidean distance, and created a hierarchical cluster tree using the nearest distance (Martinez et al., 2004). The k-means classification technique used the Euclidean distance as the distance measure and the Expectation-Maximization algorithm was applied using the Gaussian mixture model (Martinez et al., 2004). The number of clusters was set to increase from one to ten clusters. The performance of each cluster technique was measured by assessing the ability to explain variances in jump height (dependent variable) across generated clusters. This approach was based on the assumption that an appropriate grouping of subjects does not mask performance related factors and hence enhances the ability to describe variances in jump height. In order to assess the ability to explain variances in jump height for a given number of clusters the average r^2-value of a stepwise regression analysis was computed across these clusters. Input variables for the regression model were similarity scores measured solely over the key phases of a cluster. If the stepwise regression analysis was not able to identify any predictor variables within a cluster, the highest r^2-value computed during the correlation analysis (between the generated subject scores and jump height) was used. If a cluster technique assigned only one participant to a cluster, the cluster was discarded. The entire process was repeated 10 times using different random initial weights in the k-means and model based clustering to achieve a repeatable measure of the expected accuracy. The clustering technique with x groups that generated the highest stable ability to explain variances in jump height was considered the most appropriate clustering technique. To understand the underlying neuromuscular capacities of the generated clusters the following section performed a one-way ANOVA (Bonferroni adjustment for multiple comparisons) to identify differences between the generated clusters using joint angle, angular velocity, joint moment and joint power of the ankle, knee and hip joint. All statistical analyzes were performed using MatLab (R2012a, MathWorks Inc., USA).

Results: Hierarchical clustering (normalized scores) reached its highest stable ability to describe jump height using seven clusters (92 %), k-means (normalized scores) reached its highest stable ability using four clusters (90 %), k-means (non-normalized scores) reached its highest stable ability using six clusters (88 %) and hierarchical clustering (non-normalized scores) reached its highest stable ability to describe jump height using five clusters (88 %). The Expectation-Maximization algorithm did not show a stable ability to describe jump height for any number of clusters. Hierarchical clustering (normalized scores) with seven clusters was able to describe jump height best. However, varying the number of clusters in hierarchical clustering (normalized scores) between four and eight clusters had no major impact on the ability to explain jump height (1 %), while using four clusters results in greater sample sizes within the clusters. Due to the insignificantly lower ability to describe jump height and the greater sample sizes using four clusters (allowing stronger statistical analysis), the authors' decided to use hierarchical clustering (normalized scores) with four clusters for further analyses, rather than the seven clusters. Cluster 1, 2, 3 and 4 contained 6, 40, 25 and 52 subjects, respectively. Due to the small sample size in cluster 1 and the resulting limited statistical power and increased probability of committing a type II error (Cohen, 1988), cluster 1 was discarded from further analysis. The statistical analysis for

differences between the clusters 2, 3 and 4 indicated significant differences over numerous phases in joint angles, angular velocity, joint moment and joint power. Differences between cluster groups are detailed in Table 1.

Table 1: Significant differences in joint angle, angular velocity moment and power of the ankle, knee and hip joint between cluster 2, 3 and 4. The 'phase' column reports the phase over which significant differences occored and the 'differences' column reports which culsters differed from each other.

	Ankle Joint		Knee Joint		Hip Joint	
	phase	differences	phase	differences	phase	differences
Joint angle	1-25%	C3 > C2, 4	1-25 %	C3, 4 > C2	1-21 %	C4 >3> C2
	67-84%	C3 > C2, 4	67-85 %	C3, 4 > C2	65-83 %	C4 > C2, 3
	--	---	--	--		C3 >C2
	98-100%	C3 >C2, 4	99-100 %	C3,4 >C2	99-100 %	C4 >C2, 3
Angular velocity	--	--	8-19 %	C4 >C2, 3	--	--
	29-41%	C3, 4 >C2	37-50 %	C4 >C2, 3	31-41 %	C2, 3 > C4
	54-66%	C4 >C2	--	--	51-60 %	C2, 3 > C4
	--	--	66-79 %	C3 >C2	--	--
	--	--		C4 > C2	--	--
	81-89%	C4 >C2	--	--	--	--
	--	--	99-100 %	C2 >C3, 4	--	--
Joint moments	1-6%	C4 >C2	1-7 %	C2, 4 >C3	1-6 %	C3 > C4
	29-45%	C4 > C2	34-49 %	C4 >C2, 3	24-40 %	C3 > C4
	62-75%	C4 >C3	--	--	71-78 %	C2 >C4
	89-96%	C2 >C3, 4	89-95 %	C2 >C3, 4	85-89 %	C2 >C4
Joint power	--	--	1-2 %	C4 >C3	--	--
	9-19%	C4 >C2	10-28 %	C4 >C2, 3	18-31 %	C3 > C4
	32-44%	C3, 4 >C2	--	--	--	--
	57-68%	C4 >C2	50-65 %	C4 >C2	49-60 %	C2, 3 >C1, 4
	79-86%	C4 >C3	--	--	83-88 %	C2 >C4
	93-98	C2 >C3	91-96 %	C2 >C3, 4	--	--

DISCUSSION: The examined clustering techniques differed in their performance. The k-means (normalized scores) and hierarchical clustering (normalized scores) demonstrated the best abilities to identify movement strategies within the data. This highlights the use of k-means and hierarchical and the importance of normalizing subject scores when identifying movement strategies. Normalizing subject scores (transformation of scores into their correlation matrix) had a significantly positive effect on the performance of both hierarchical and partitional clustering techniques, indicating that differences in magnitude between subject scores are not as effective as their quantified numerical relationship at maximising the ability to identify movement strategies. Non-normalized scores describe similarity in movement strategy using a combination of the magnitudes within each of the identified phases of variation. Hence, different movements strategies with similar magnitudes (e.g. in knee moment) may erroneously appear more similar than movement strategies with different magnitudes. In contrast, normalized scores describe similarity of a movement based on the pattern in the magnitudes between the identified phases.

The optimal clustering approach generated 3 distinct movement strategies. Defining characteristics of cluster 2 are low knee and hip joint angles (greater joint flexion), the ability to generate large knee moments, to maintain hip moments towards the end of the movement cycle and a delayed ankle, knee and hip peak power. Defining characteristics of cluster 3 are high ankle and knee joint angles (greater joint extension) throughout the movement cycle, the inability to generate large ankle and knee moments, the ability to generate large initial hip moments and the inability to maintain large moments towards the end of the movement cycle. Defining characteristics of cluster 4 are high ankle moment throughout much of the movement cycle, the ability to generate large initial knee moments, and the inability to generate large hip moments. Given that there are a number of movement strategies, this may explain the inconsistencies in previous findings. The number of participants with particular strategies (e.g. cluster 2, 3 and 4) can influence the finding within a group analysis design. For example, some studies found peak moments in the knee joint to be greater than the ankle joint (Bobbert, Huijing, & van Ingen Schenau, 1987; Vanrenterghem, Lees, & De Clercq, 2008), while others found the opposite (AragonVargas & Gross, 1997; Vanezis & Lees, 2005). The number of participants with particular strategies may also explain, at least in part, why performance related factors identified across studies differed in previous studies. This highlights the use of a subgroup analysis.

CONCLUSION: Hierarchical clustering utilizing normalized subject scores to generate 4 clusters appears to be the most suitable technique for clustering force curves, while k-means clustering (normalized subject scores with 4 clusters) also showed a high level of suitability. The generated clusters demonstrated clear defining characteristics, which at least in part explain inconsistencies in findings of previous studies. Consequently, utilizing a subgroup analysis might give a better insight into which factors relate to performance in a movement task. This allows the optimization of a training interventon in relation to a specific movement strategy.

REFERENCES

AragonVargas, L. F., & Gross, M. M. (1997). Kinesiological Factors in Vertical Jump Performance : Differences Among Individuals. Journal of Applied Biomechanics, 13(1), 24–44.
Bobbert, M. F., Huijing, P. A., & van IngenSchenau, G. J. (1987). Drop jumping. I. The influence of jumping technique on the biomechanics of jumping. Medicine and Science in Sport and Exercise, 19(4), 332–338.
Cohen, J. (1988). Statistical power analysis for the behavioralsciencies.Routledge.
Jain, A., Murty, M., & Flynn, P. (1999). Data clustering: a review. ACM Computing Surveys (CSUR), 31(3), 264–323.
Martinez, W. L., Martinez, A., &Solka, J. (2004). Exploratory data analysis with MATLAB.CRC Press.
Richter, C., O'Connor, N. E., Marshall, B., & Moran, K. (2013a). Analysis of Characterizing Phases on Waveforms – An Application to Vertical Jumps.Journal of Applied Biomechanics (in Press).
Richter, C., McGuiness, K., Gualano, L., O'Connor, N. E., & Moran, K. (2013b). Identification of an optimal principal components analysis threshold to describe jump height accurately using vertical ground reaction forces.In XXIV Congress of the International Society of Biomechanics. Natal, Brazil.
Stergiou, N. (2004). Innovative Analyses of Human Movement (1st ed.). Leeds, U.K.: Human Kinetics.
Stergiou, N., & Scott, M. (2005). Baseline measures are altered in biomechanical studies. Journal of Biomechanics, 38(1), 175–178.
Toro, B., Nester, C. J., &Farren, P. C. (2007). Cluster analysis for the extraction of sagittal gait patterns in children with cerebral palsy. Gait & Posture, 25(2), 157–165.
Vanezis, A., & Lees, A. (2005). A biomechanical analysis of good and poor performers of the vertical jump. Ergonomics, 48(11-14), 1594–603. doi:10.1080/00140130500101262
Vanrenterghem, J., Lees, A., & De Clercq, D. (2008). Effect of forward trunk inclination on joint power output in vertical jumping. The Journal of Strength & Conditioning Research, 22(3), 708–714.
Winter, D. A. (2009). Biomechanics and motor control of human movement.Wiley.com.