Clinical Relevance of Grouping Runners with Knee Pain through Statistical Clustering Analyses

Madeline Grosklos[1,2], Adam Bunn[1,2], Marisa Pontillo[1,2]
[1]Extremity Trauma and Amputation Center of Excellence, San Diego, CA, USA
[2]Naval Medical Center San Diego, San Diego, CA, USA
Email: grosklos.2@osu.edu

Summary

Knee pain is common in runners and a cause for limited duty in the military. Quantitative methods of classifying movement in individuals with knee pain could improve run retraining strategies. This study utilized statistical parametric mapping to understand whether clinically meaningful differences were present in clusters of runners with knee pain formed through principal component analysis and Dirichlet-Process K-means clustering (DP-means). The clustering methods successfully separated runners with more extended-leg running from those with greater amounts of joint flexion. Differences in sagittal plane moments, gait speed, and post-running session pain, were also apparent.

Introduction

In 2005, lower extremity overuse injuries resulted in an estimated 3,800,000 limited duty days in U.S. military Service members [1] and in 2006, the knee/lower leg comprised the largest proportion of all injury-related musculoskeletal conditions [2]. As these injuries are caused by repetitive strain, one treatment approach lies in retraining movement strategies during cyclic activities such as running. This study investigated whether groups of runners with knee pain, formed through numerical methods, displayed clinically distinct movement patterns that could be used to refine movement retraining in the future.

Methods

50 active-duty service members with unilateral knee pain (35M/15F; 32.3±8.0 years) underwent 3D motion capture of running at a self-selected speed on an instrumented treadmill. Knee pain (visual analog scale) was also collected. A sex-specific model with functional hip and knee joint centers was applied to each participant in Visual3D. Kinematic, kinetic, and spatiotemporal data were averaged over one 20-second trial. Multivariate functional principal component analysis, run with sagittal plane hip/knee moments, sagittal/frontal plane hip/knee/pelvis angles, and sagittal ankle angles, and DP-means clustering were performed in R (v4.4.2); 46 of the participants fell within two clusters (C1 and C2). Differences in running biomechanics between the clusters were assessed through statistical parametric mapping general linear models run in MATLAB (2024a, spm1d vM.0.4.10) with gait speed as a covariate [3]. Demographics, patient reported outcomes, and spatiotemporal gait parameters were compared using Mann-Whitney U tests in SPSS statistics (v29).

Results and Discussion

C1 and C2 contained 27 (9 female) and 19 (6 female) participants, respectively. After controlling for gait speed, C2 demonstrated less anterior pelvic tilt, less hip, knee, and ankle flexion/dorsiflexion, greater knee adduction angle, and smaller external hip, knee, and ankle flexion/dorsiflexion moments during the stance phase of running on the involved limb compared to C1. For the uninvolved limb, knee flexion moment, knee adduction angle, and ankle dorsiflexion moment did not differ (Table 1). C2 had a slower running speed [2.31 (.66)m/s, 2.67 (.58)m/s], shorter stride length (normalized to height) [.97 (.18), 1.17 (.24)], longer stance time [.40 (.07)s, .36 (.06)s $p \leq .036$], and higher pain after the session [4 (3),1.4 (4)] than C1 ($p \leq .032$). C2 had smaller flexion angles at the hip, knee, and ankle independent of running speed, indicating a stiffer, more extended lower-limb posture during stance compared to C1. Lower flexion angles and moments could be a compensatory mechanism unique to participants in C2 that impairs energy absorption and increases risk of stress injuries to bones [4]. This could partially explain why C2 reported higher pain than C1 immediately after the running session, but not at the time the session started.

Conclusions

The PCA and clustering analysis successfully separated participants with unique running patterns and the differences were interpretable with clinical relevance. Such methods of clustering runners may aid in determination of movement retraining approaches for each participant.

Acknowledgments

The views expressed herein are those of the author(s) and do not reflect the official policy or position of the Naval Medical Center San Diego, the Defense Health Agency, the Department of Defense, or any agencies under the U.S. Government. This project was supported in part by an appointment to the Research Participation Program for the Sponsor administered by the Oak Ridge Institute for Science and Education through an agreement between the U.S. Department of Energy and the Agency.

References

[1] Jones BH et al. (2010). *AJPM*, **38**: S42-S60.
[2] Hauret KG et al. (2010) *AJPM,* **38**: S61-S70.
[3] Pataky TC (2010). *J. Biomech.*, **43**: 1976-1982.
[4] Warden SJ et al. (2014). *JOSPT,* **44**: 749-765.

Table 1: SPM results. The region indicates the percentage of stance phase where differences (C1 > C2) were observed. ND = no difference.

		Angles					Moments		
Side	SPM	Pelvis X	Hip X	Knee X	Knee Y	Ankle X	Hip X	Knee X	Ankle X
Painful	p-value	<.001	<.001	<.001	.047	<.001	<.001	<.001	.002
	Region	0-100%	0-100%	7-93%	66-72%	12-89%	75-98%	22-71%	64-88%
Non-painful	p-value	<.001	<.001	.001	ND	<.001	<.001	ND	ND
	Region	0-100%	0-100%	24-65%		23-81%	75-100%		