A Data Mining Approach for Determining Gait Abnormalities in Runners with Patellofemoral Pain Syndrome

Ross Brancati, MS

Co-author: Katherine Boyer, PhD

University of Massachusetts Amherst

Department of Kinesiology

Musculoskeletal and Orthopedic Biomechanics Laboratory

@ross_brancati

/ross-brancati/

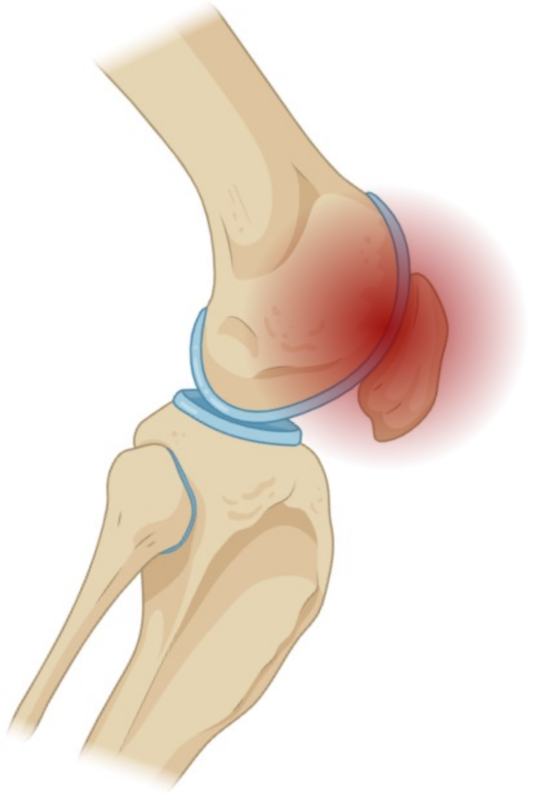
rbrancati@umass.edu

rossbrancati.github.io

NACOB
NORTH AMERICAN CONGRESS ON BIOMECHANICS
OTTAWA 2022

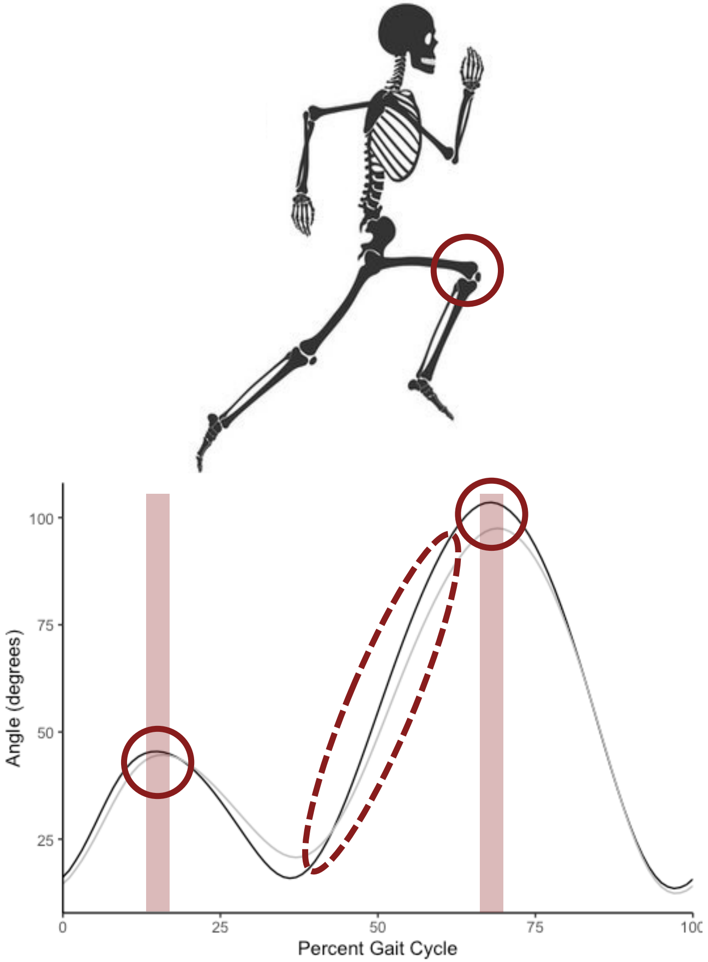
Patellofemoral Pain Syndrome

- Patellofemoral pain (PFP) is a common musculoskeletal pain disorder, often presenting as anterior knee pain
- Although PFP occurs at the knee, distal and proximal joints should be considered as well
- Data mining approaches may offer better insight into the complex etiology of PFP

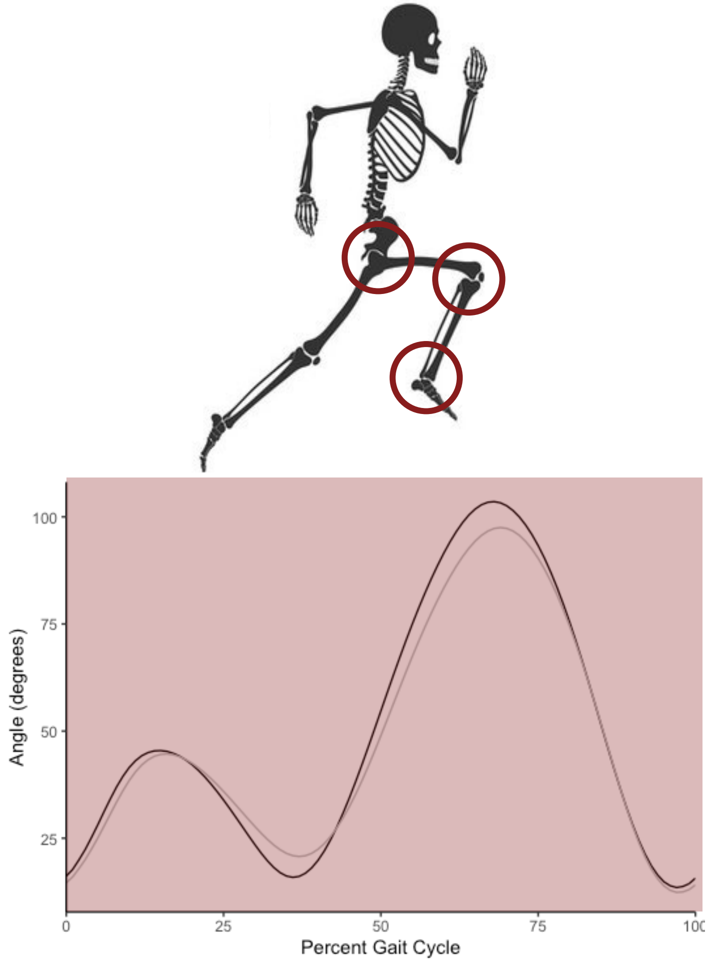


Taunton et al. 2002, Powers et al. 2017, Willy et al. 2019

Prior Work



Isolated joints and discrete variables



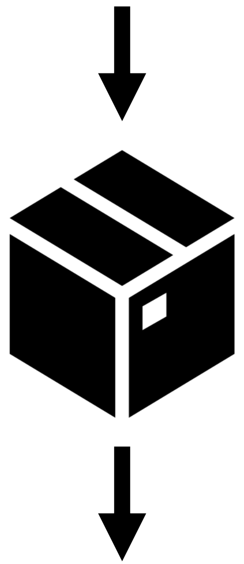
Multiple joints and timepoints of the gait cycle

Barton et al. 2012, Dierks et al. 2011, Esculier et al. 2015, Noehren et al. 2012

Interpretable Machine Learning (ML)

Black Box Approaches

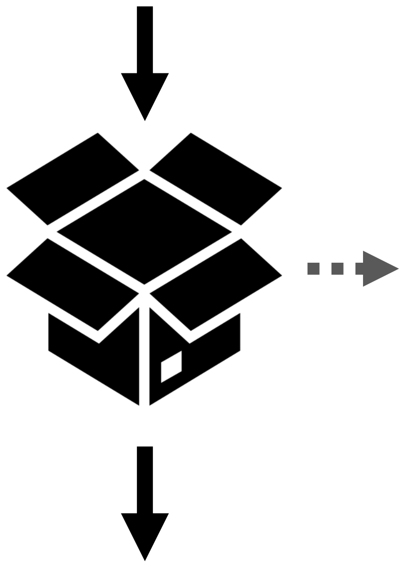
Inputs



Outputs

Interpretable ML Approaches

Inputs



- Informativeness
- Transferability
- Explainability
- Feature Importance

Outputs

Methods – Data Collection

41 recreational runners
- 15 healthy (H)
- 14 symptomatic PFP (S)

Runners completed a 21-minute treadmill run at self-selected speed

- M1 ——— 21-minute treadmill run ——— M21 -

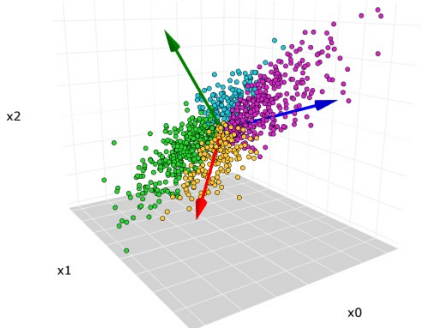
Kinetic, kinematic, electromyography, RPE, and subjective pain scores collected at 1st (M1) and 21st (M21) minutes

Methods – Data Processing

1. *Generate variable matrices*

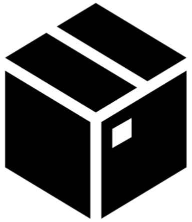
Subject ID	Group	t1	t2	…	t101
Subject 1 M1	H	Datapoint	Datapoint	…	Datapoint
Subject 1 M21	H	Datapoint	Datapoint	…	Datapoint
Subject 2 M1	S	Datapoint	Datapoint	…	Datapoint
…	…	Datapoint	Datapoint	…	Datapoint
Subject 29 M21	R	Datapoint	Datapoint	…	Datapoint

2. *Principal Component Analysis*



→ Generates Principal Component (PC) Scores

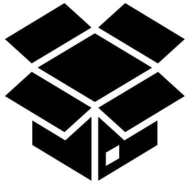
3. *Logistic Regression (LR) Classifier*

PC Scores →  → Predictions

LR Classifier

4. *Feature Selection*

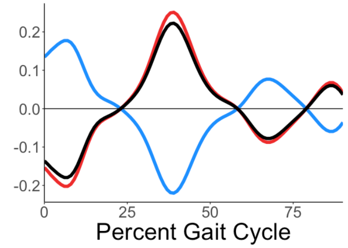
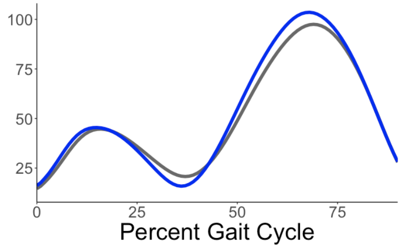
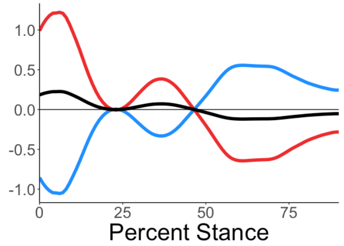
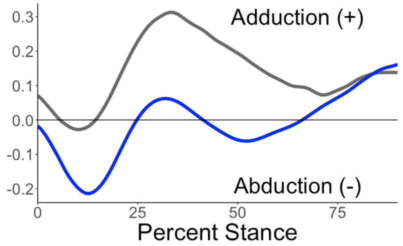
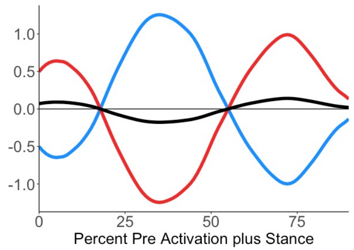
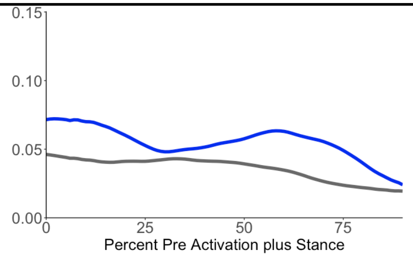
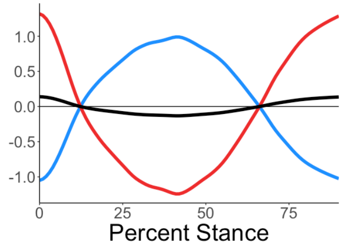
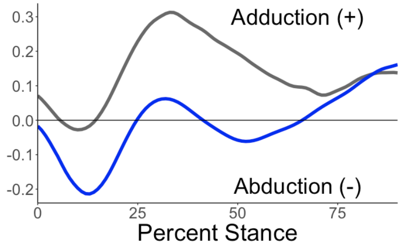
$$h_\theta(x_i) = g(\theta_0 + \sum_{i=1}^{m} \boldsymbol{\theta_i} x_i)$$

 → $\boldsymbol{\theta_1, \theta_2, \dots, \theta_n}$

Extract Feature Coefficients

Results – Minute 1

Model Accuracy: 82.76%

Feature	PC Plot	Waveforms	Logistic Regression Coefficient	PC Interpretation
Knee Flexion Angle PC3			3.617	Related to peak stance angle, toe-off angle, and peak swing angle
Hip Adduction Moment PC3			1.301	Related to peak hip abduction moment in early and late stance
Glute Max Activation PC2			0.929	Related to magnitude of activation in pre-stance and late stance
Hip Adduction Moment PC1			0.801	Related to overall magnitude of moment in mid-stance

PC

5th %ile

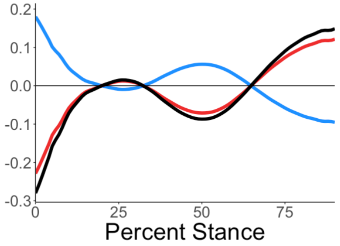
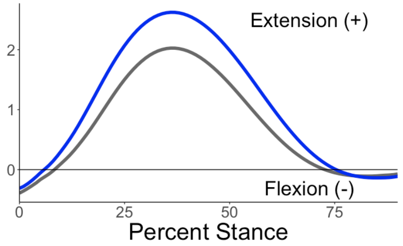
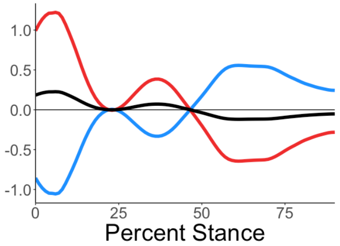
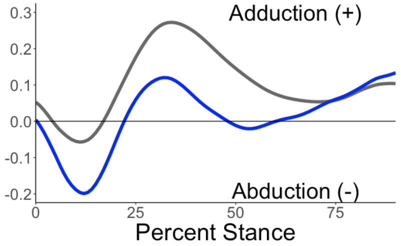
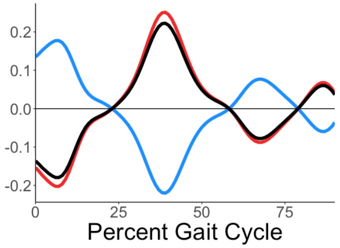
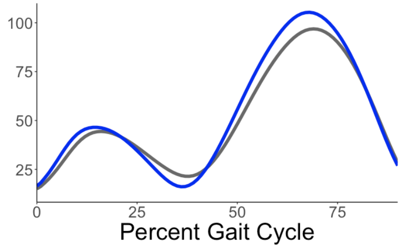
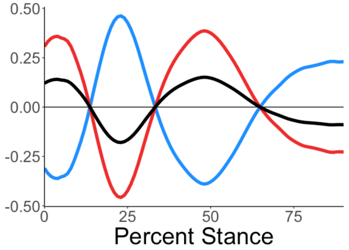
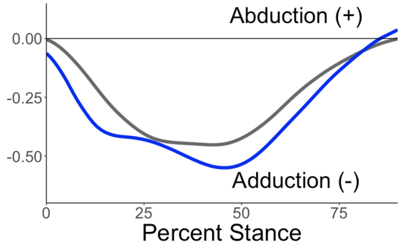
95th %ile

Healthy

Injured

Results – Minute 21

Model Accuracy: 79.31%

<table>
<tr><th>Feature</th><th>PC Plot</th><th>Waveforms</th><th>Logistic Regression Coefficient</th><th>PC Interpretation</th></tr>
<tr><td>Knee Flexion Moment PC2</td><td>
</td><td>
</td><td>2.341</td><td>Related to loading rates in early and late stance</td></tr>
<tr><td>Hip Adduction Moment PC3</td><td>
</td><td>
</td><td>1.608</td><td>Related to peak hip abduction moment in early and late stance</td></tr>
<tr><td>Knee Flexion Angle PC3</td><td>
</td><td>
</td><td>1.515</td><td>Related to peak stance angle, toe-off angle, and peak swing angle</td></tr>
<tr><td>Knee Adduction Moment PC3</td><td>
</td><td>
</td><td>1.479</td><td>Related to 1st and 2nd peak moments, or magnitude of moment in stance</td></tr>
</table>

PC

5th %ile

95th %ile

Healthy

Injured

UMass MOBL
Musculoskeletal Orthopedic Biomechanics Lab

University of Massachusetts Amherst
BE REVOLUTIONARY

Results - Summary

Group	Pain Level Change (M1 > M21)	BORG RPE Change (M1 > M21)
Healthy	0	5
Symptomatic	3	3

At M1, Healthy group demonstrated

- Greater range of knee flexion
- Larger magnitude of hip abduction loading
- More glute max activation

At M21, Healthy group demonstrated

- Greater loading in knee flexion
- Larger magnitude of knee adduction loading

Conclusions

Dimensionality reduction and interpretable machine learning approaches capture important biomechanical adaptations.

Sagittal plane knee angles and frontal plane hip angles are pertinent features in runners with PFP.

With the onset of pain, runners with PFP avoid knee flexion and adduction loading patterns.

Future research should consider systematic approaches to optimize interventions in clinical populations.

Acknowledgements

MOBL Lab:

- Dr. Katherine Boyer
- Kiichi Ash
- Erica Casto
- Skylar Holmes
- Aidan Gross
- Kali Shamaly
- Athulya Simon
- Carl Jewell

DE LUCA FOUNDATION

Funding from the De Luca Foundation

NACOB
NORTH AMERICAN CONGRESS ON BIOMECHANICS
OTTAWA 2022

ASB
American Society of Biomechanics

Canadian Society for Biomechanics
Société canadienne de biomécanique

@ross_brancati
rbrancati@umass.edu
/ross-brancati/
rossbrancati.github.io

Questions?

References

Taunton, J. E., Ryan, M. B., Clement, D. B., McKenzie, D. C., Lloyd-Smith, D. R., & Zumbo, B. D. (2002). A retrospective case-control analysis of 2002 running injuries. *British Journal of Sports Medicine*, *36*(2), 95–101. https://doi.org/10.1136/bjsm.36.2.95

Powers, C. M., Witvrouw, E., Davis, I. S., & Crossley, K. M. (2017). Evidence-based framework for a pathomechanical model of patellofemoral pain: 2017 patellofemoral pain consensus statement from the 4th International Patellofemoral Pain Research Retreat, Manchester, UK: Part 3. *British Journal of Sports Medicine*, *51*(24), 1713–1723. https://doi.org/10.1136/bjsports-2017-098717

Barton, C. J., Levinger, P., Crossley, K. M., Webster, K. E., & Menz, H. B. (2012). The relationship between rearfoot, tibial and hip kinematics in individuals with patellofemoral pain syndrome. *Clinical Biomechanics*, *27*(7), 702–705. https://doi.org/10.1016/j.clinbiomech.2012.02.007

Dierks, T. A., Manal, K. T., Hamill, J., & Davis, I. (2011). Lower extremity kinematics in runners with patellofemoral pain during a prolonged run. *Medicine and Science in Sports and Exercise*, *43*(4), 693–700. https://doi.org/10.1249/MSS.0b013e3181f744f5

Esculier, J. F., Roy, J. S., & Bouyer, L. J. (2015). Lower limb control and strength in runners with and without patellofemoral pain syndrome. *Gait and Posture*, *41*(3), 813–819. https://doi.org/10.1016/j.gaitpost.2015.02.020

Noehren, B., Sanchez, Z., Cunningham, T., & McKeon, P. O. (2012). The effect of pain on hip and knee kinematics during running in females with chronic patellofemoral pain. *Gait and Posture*, *36*(3), 596–599. https://doi.org/10.1016/j.gaitpost.2012.05.023

Molnar, C. (2019). Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. *Leanpub Publishing*.

Halilaj, E., Rajagopal, A., Fiterau, M., Hicks, J. L., Hastie, T. J., & Delp, S. L. (2018). Machine learning in human movement biomechanics: Best practices, common pitfalls, and new opportunities. In *Journal of Biomechanics* (Vol. 81, pp. 1–11). Elsevier Ltd. https://doi.org/10.1016/j.jbiomech.2018.09.009

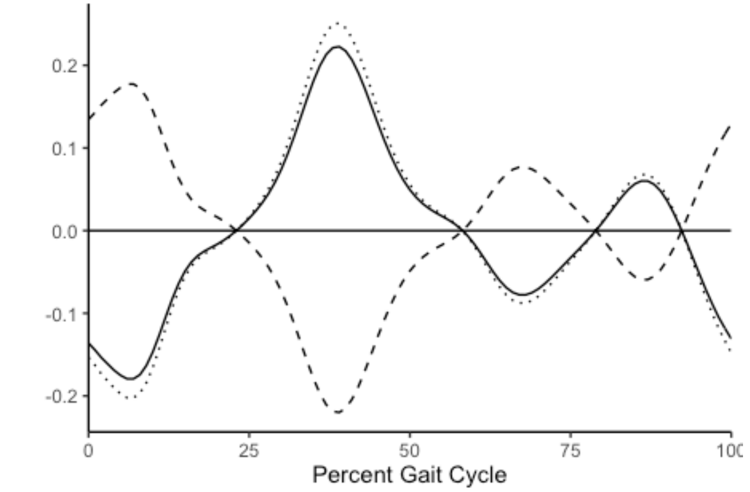
Deluzio, K. J., & Astephen, J. L. (2007). Biomechanical features of gait waveform data associated with knee osteoarthritis. An application of principal component analysis. *Gait and Posture*, *25*(1), 86–93. https://doi.org/10.1016/j.gaitpost.2006.01.007

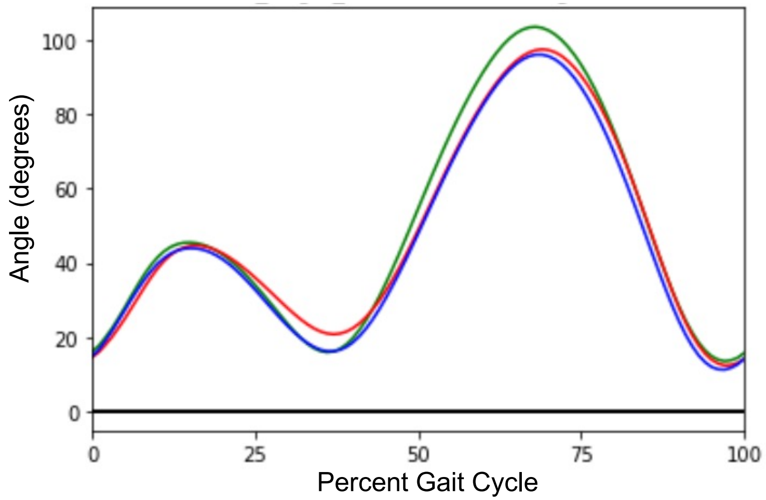
Cheng, C. (2022) Principal Component Analysis (PCA) Explained Visually with Zero Math. *Towards Data Science.* n.d. Accessed: 8/2/2022

Slide 2 image credit: Adobe Stock. Accessed 8/2/2022 https://stock.adobe.com/ee/search?k=human+skeleton

Interpreting Principal Components

Example Principal Component (PC): Knee Flexion Angle PC 3





Principal Component (aka eigenvector)

5th p

95th

Healthy g
Symptom
Recovere

Interpretation: a more negative score (more similar to the healthy group) indicated greater peak angle during stance, lesser angle at toe off, and greater peak angle during swing.

Hea

Symptomatic group ensemble average

Recovered group ensemble average

Interpretation method from Deluzio et al. 2007

UMass MOBL Musculoskeletal Orthopedic Biomechanics Lab

University of Massachusetts Amherst BE REVOLUTIONARY