Unsupervised Clustering of Acoustic Emissions to Explore Correlations with Patient-Reported Outcomes.

H. E. L. Rose[1], D. Tsaopoulos[2], S. Nikolopoulos[2], **D-S. Komaris[2]**
[1]Medical Physics and Clinical Engineering, Nottingham University Hospitals NHS Trust, Nottingham, UK
[2]Centre for Research & Technology Hellas, Information Technologies Institute, Thessaloniki Greece.
Email: s.komaris@iti.gr

Summary

This study explores using PCA and unsupervised clustering to classify knee joint acoustic emissions (AEs) and identify early joint degeneration. Data from fifty adults were used, with AE metrics captured during cycling. Gaussian Mixture Models revealed two clusters differing in AE features, PROMs (5STS), and subject age. Findings highlight AE's potential clinical value.

Introduction

Balancing objective clinical assessments, like imaging, with patient-reported outcomes (PROMs) is crucial for personalized medicine but remains challenging due to costs, accessibility, and variability in PROMs influenced by psychological factors. Acoustic Emission (AE) monitoring, a non-invasive diagnostic tool, has emerged as a dynamic alternative, capturing real-time sound waves from movements to assess joint integrity, with the potential to detect early joint degeneration [**1**]. Machine learning enhances AE data interpretation but faces challenges in separating healthy from pathological knees due to potential biases. This study uses unsupervised clustering on a diverse cohort to explore relationships between AE-derived clusters and PROMs.

Methods

A dataset [**2**] containing AE records of 50 younger adults (18–35 years, 13 males and 12 females) and older adults (50–75 years, 14 males and 11 females) was used for clustering analysis. Detailed participant demographics and acquisition procedures are provided in the original dataset publication [**2**]. Knee AE recordings were conducted during a 1-minute cycling trial at 60 rpm using a USB AE Node system with a sampling rate of 20 Msps and a 20 dB gain (PK3I sensor), capturing features including acoustic event (hit) amplitude, duration (μs), absolute energy (aj), signal strength (pV-s), rise time (μs), and hit count (n). AE metrics were normalized between 0 and 1. The Five Times Sit-to-Stand Test (5STS) was used as a PROM [**3**]. The final analysis included 9,064 events across all 50 participants. Unsupervised clustering with Gaussian Mixture Models (GMMs) was performed to identify patterns in AE metrics. Optimal cluster numbers (2–10) were determined using Silhouette Scores, and Bayesian Information Criterion (BIC) guided model selection across 14 geometry configurations. Differences in AE metrics, PROMs, and demographics were analysed between clusters using non-parametric Wilcoxon rank-sum tests.

Results and Discussion

PCA identified two principal components (PCs) explaining 97% of the dataset's variance. PC1 was mainly influenced by rise time and duration, while duration and hit count dominated PC2. A GMM with two clusters and an ellipsoidal geometry (VEE model) performed best, achieving an average Silhouette score of 0.35 and a BIC of 160301.4. Scatter plots of PC1 and PC2 for the clusters are shown in **Figure 1**, while **Table 1** compares cluster means. Significant differences ($p < 0.0036$, Bonferroni-corrected) were observed across all AE metrics in **Table 1**, as well as subject age and 5STS times.

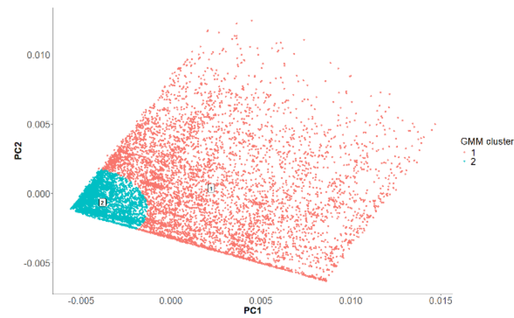

Figure 1: Scatter plots of PC1 (●) and PC2 (●) for the GMM.

Table 1 Means and Statistical Comparisons Between GMM Clusters.

Metric	Cluster 1	Cluster 2	Full Cohort
Absolute Energy, aJ	1.74e-06	5.83e-07	1.29e-06
Counts	3.67e-03	1.46e-03	2.80e-03
Duration, μs	5.04e-03	1.60e-03	3.69e-03
Rise time, μs	6.16e-03	1.25e-03	4.23e-03
Signal Strength, pV-s	3.90e-04	1.26e-04	2.87e-04
5STS (seconds)	8.10	7.34	7.80
Subject Age (years)	47	44	46

Conclusions

This study highlights the feasibility of using unsupervised GMM and PCA to classify knee joint AEs. Challenges in optimizing cluster numbers can lead to artificial groupings, with clusters without clear clinical relevance. The association between AE features and PROMS (5STS, **Table 1**) supports the clinical value of the present clustering approach. Notably, Cluster 1 was characterized by older participants, further underscoring the potential relationship between AE patterns and participant demographics [**4**].

References

[1] Khan T. et al. (2024). Bio. Sign Proc. Control, **87**: 105510
[2] Khokhlova L. et al. (2024). PLoS ONE, **19**: e0310123
[3] Whitney S. et al. (2005). Physical therapy, 10: 1034-45.
[4] Shark, L-K et al. (2011). Med. Eng. & phys., 33: 534-45.