Predicting Phenotypes of Osteoarthritis Progression:
Eight-Year Data from the Osteoarthritis Initiative

Eni Halilaj, Ya Le, Jennifer L. Hicks, Trevor J. Hastie, Scott L. Delp
Stanford University, Stanford, CA, USA
Mobility is essential for human health

Decreased physical activity is linked to heart disease, stroke, diabetes, depression, and obesity.
Mobilize Center at Stanford University

PI: Scott Delp

Our mission is to overcome the data science challenges facing biomedical big data to improve human movement across the wide range of conditions that limit mobility.
Discovering Insights to Improve Mobility

**Treatment Planning with Biomechanics & Data Science**

**Motivating Physical Activity at a Global Scale**

**Integrating Multimodal Data to Improve Rehabilitation**
Knee Osteoarthritis (OA)

• The most common cause of limited mobility in older adults¹
• One in two individuals are at risk of developing knee OA by age 85²
• Prevalence is projected to rise by 50% by 2040³
• Current treatments are limited to symptoms management
• One of our goals is to characterize and predict disease progression phenotypes in order to improve clinical trial design

¹ Guccione et al., 1994  ² Murphy et al., 2008  ³ Hootman et al. 2016
Eight-year study of the natural progression of knee OA whose mission is to accelerate the development of treatments
**Study Goals**

1. Determine the most prominent phenotypes of osteoarthritis progression by clustering eight-year disease progression data.

![Graphs showing disease progression over time with k = 1, k = 2, and k = 3 clusters.]

2. Identify features available at the enrollment visit that are predictive of disease progression phenotype.
Structural Progression: Joint Space Width

Medial

Lateral

Symptomatic Progression: Pain

**QUESTION:** How much pain have you had . . .

1. when walking on a flat surface?
   - none
   - mild
   - moderate
   - severe
   - extreme

2. when going up or down stairs?
   - none
   - mild
   - moderate
   - severe
   - extreme

3. at night while in bed? (that is - pain that disturbs your sleep)
   - none
   - mild
   - moderate
   - severe
   - extreme

4. . . .

Clustering of Functional Data

Model progression in terms of time and use an iterative procedure to solve for the model parameters (e.g., cluster means) that maximize the likelihood of the observed data.
Clustering of Sparse Functional Data

mixed-effects mixture model

Profile 1: Worsens, then stabilizes

Profile 2: Worsens continuously

Model progression in terms of time and use an iterative procedure to solve for the model parameters (e.g., cluster means, covariance structure) that maximize the likelihood of the observed data

James and Sugar, 2003
Joint Space Narrowing

n = 973

n = 270
Pain Progression

Change in pain score (0 – 20)

Years since enrollment

Mean

95% CI

Joint space narrowing (mm)

n = 973

0           1           2          3           4           5           6          7

0.0

-0.5

-1.0

-1.5

-2.0

-2.5

n = 270

Mean

95% CI
Predicting Progression Phenotypes

Gradient Boosted Trees Model

AUC = .63

Variables and their relative variable importance:
- Max Adult Weight
- Limb Malalignment
- History of Osteoporosis
- Years of Smoking
- Daily Fat Consumption
- Blood Pressure
- Sex
Physical Activity and Cartilage Health

Univariate Approach

R = .09, p = .71

Mean T2 Signal Δ (ms)

Total Amount of Activity (counts)
Physical Activity and Cartilage Health

Univariate Approach

Total Amount of Activity (counts)

Mean T2 Signal Δ (ms)

R = .09, p = .71
R = .98, p = .00

Multivariate Pattern Analysis

Linear Combination of Activity Types

Linear Combination
of T2 Δ in Different Regions

R = .98, p = .00
Summary

**Key Finding 1:** Structural knee OA progresses very slowly, but there is also a fast-progressing phenotype

**Key Finding 2:** Pain and structural progression are not concomitant

**Key Finding 3:** OA patients should be monitored and classified as fast or slow progressors using predictive models before enrollment in clinical trials

**Future Work:** Incorporate more sensitive imaging (MRI) and biospecimen markers, as well as physical activity data from mobile sensors in predictive models
Growing the Biomedical Data Science Community

The Mobilize Center is funded by NIH Grant U54EB020405