THE MOBILIZE CENTER: ACCELERATING MOVEMENT SCIENCE WITH BIG DATA

Eni Halilaj[1], Jennifer L. Hicks[1], Joy P. Ku[1], Scott L. Delp[1,2]
Departments of Bioengineering[1] and Mechanical Engineering[2], Stanford University
Corresponding author email: delp@stanford.edu

INTRODUCTION

Mobility is essential for human health. Unfortunately, many musculoskeletal conditions, including cerebral palsy, osteoarthritis, running injuries, obesity, and stroke, limit mobility at a great cost to quality of life and healthy aging. The proliferation of wearable devices that monitor human activity is generating unprecedented quantities of data on human movement, behaviors, and health. Mobility data is also being collected daily by hundreds of clinical centers and research laboratories around the world. The mission of the Mobilize Center (mobilize.stanford.edu), an NIH Big Data to Knowledge (BD2K) Center of Excellence, is to overcome the data science challenges posed by mobility big data and improve human movement across the wide range of conditions that limit mobility. In this talk we will provide an overview of the Center's research, training, and dissemination activities and share specific ways in which the biomechanics community can become engaged.

METHODS

Our data science efforts center around four themes. We are developing (1) robust and flexible optimization tools to generate personalized biomechanical models and simulations from diverse experimental movement data, including wearable sensors; (2) new statistical learning algorithms to make predictions and discover patterns from large sets of noisy, sparse, and complex data, whether discrete or time-varying; (3) tools to model the role of behavioral and social dynamics in human health based on information collected with smartphones and wearable activity monitors; and finally, (4) machine-learning systems that integrate unlabeled data from diverse sources to aid clinical decision-making and transparently communicate with clinicians (Figure 1). We are disseminating these tools as open-source packages.

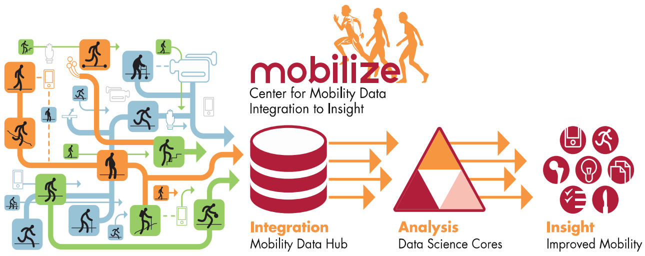

Figure 1: **Vision of the Mobilize Center.** We have partnered with hospitals, biomechanics labs, and industry affiliates to assemble a massive database describing human motions, including trajectories of markers placed on the body, video, ground forces, range-of-motion measurements, muscle electromyography, muscle strength, accelerometer and GPS recordings from wearable sensors, food intake, sleep records, and electronic health records. Analyses of mobility big data using our tools are generating important new insights for surgical planning, gait modification, prosthesis design, and exercise prescriptions.

To ensure that our data science research has a significant impact on human health, we are focusing our activities on a few critical biomedical problems. First, we are analyzing mobility data collected at Gillette Children's Specialty Healthcare to predict and improve the outcomes of surgeries in children with cerebral palsy and gait pathology. Second, we are integrating data from biomechanics labs, hospitals, and observational cohort studies to identify new approaches to optimize mobility in individuals with osteoarthritis, stroke, running injuries, and other movement impairments. Third, we are analyzing wearable sensor data from millions of people to discover factors that motivate individuals to be more active. Fourth, we are integrating omics and portable biosensor data to enable early diagnosis of pre-diabetes through convenient mobile health platforms.

In addition to the research, the Center is also training scientists at the intersection of data science and biomechanics. Our Massive Open Online Courses, including Mining Massive Datasets, Statistical Learning, and Convex Optimization, train tens of thousands of students and researchers. We have also established a Distinguished Postdoctoral Fellows and graduate student research program to create leaders in biomedical big data analytics.

RESULTS AND DISCUSSION

Our initial analyses of large datasets have already led to new insights that could inform public policy and clinical decision-making. For example, by mining smartphone data from millions of users, we have identified new social and environmental determinants of physical activity, which we are using to design evidence-based interventions aimed at increasing activity. We have built models to predict the pace of osteoarthritis progression over the course of eight years based on data collected in one visit and are starting to better understand how habitual physical activity affects cartilage microstructure. We have also identified key features that make children with cerebral palsy good candidates for surgery.

CONCLUSIONS

The proliferation of wearable technology and big data is poised to revolutionize movement research. In the face of growing volumes of data from varying sources, the biomechanics community must come together to discuss the challenges and opportunities that come with biomechanics big data and how proper application of statistical and machine learning methods, as well as sharing of data and tools, can help us overcome some of these challenges and lead to new insights. We hope that this session and similar future efforts can serve to accelerate research at the intersection of biomechanics and data science.

ACKNOWLEDGEMENTS

The Center is supported by NIH Grant U54EB020405.

LONG-TERM MONITORING OF LOWER LIMB JOINT LOADS USING WEARABLE SENSORS: APPLICATION IN SPORT AND ORTHOPAEDICS

[1]**Thor Besier**, [2]Kaitlyn Weiss, [1]Alex Wilson, [1]Shasha Yeung, [2]Chris Whatman, [2]Michael McGuigan and [1]Justin Fernandez
[1]Auckland Bioengineering Institute, University of Auckland, Auckland, NZ
[2]Sports Performance Research Institute New Zealand, Auckland University of Technology, Auckland, NZ
Corresponding author email: t.besier@auckland.ac.nz

INTRODUCTION

Mechanical loads play a critical role in the development, maintenance, and regeneration of musculoskeletal tissue. Knowledge of musculoskeletal tissue loads during activities of everyday living or during sport is critical to prescribe appropriate interventions or identify mechanisms of injury. Traditional methods to estimate internal loads, such as muscle and joint forces, involve motion capture, force measurement, and biomechanical modelling. Although powerful, these methods are typically constrained to lab environments, thus limiting the ecological validity of the results.

Wearable sensors, such as inertial measurement units (IMUs), provide an opportunity to measure and monitor human movement over an extended period of time in real-world settings. When attached to body segments, IMUs provide useful measurement of segmental linear accelerations and angular velocities, which can be interpreted on their own, or combined with a biomechanical model to derive joint kinematics.

This paper presents a framework that integrates data from body worn IMUs with biomechanical and statistical models to measure and monitor lower limb loads. Two applications of this framework include impact load monitoring for basketball players and knee load monitoring for patients who have experience total knee joint replacement.

METHODS

The framework we have developed consists of four components and uses accelerometer and gyroscope data sampled at 1000Hz from synchronised IMUs (IMeasureU Ltd, NZ) placed on the anterior distal third of each tibia.

1. Time Series Analysis. The first step when processing large IMU data files is to segment or classify the time-series data to obtain the epoch's of interest. This step can be performed using principal component analysis (PCA) or frequency-domain methods. For example, we might pull out regions of data during the day that correspond to walking and running if we wish to analyse those separately.

2. Mechanistic Model. Prior to data collection, we perform a series of calibration tasks with the purpose of developing a surrogate model to be used in the field. We collect motion capture and ground reaction force data and perform standard inverse dynamic analysis using an OpenSim biomechanical model (Stanford, CA). Our model is scaled to match the anthropometry of the subject using the Musculoskeletal Atlas Project [1]. We collect simultaneous IMU data during these trials for the purpose of improving the time series analysis (above) and developing the surrogate model.

3. Surrogate Model. Using the IMU data from the calibration trials in conjunction with the output variables from the mechanistic model (such as joint contact force) we create a surrogate of the mechanistic model. This is merely a statistical representation that mimics the response of the mechanistic model, given IMU data as input. To achieve this we use a Partial Least Squares Regression (PLSR), which is an efficient, multivariate-fitting method. This surrogate model can then be used to predict complex output variables, such as joint loads, from field-based IMU measurements.

4. Mechanobiology model. Finally, we combine knowledge of the tissue-level output variable with load frequency using a variance of the Daily Load Stimulus (DLS), which accounts for multiple loading events of different magnitudes [2].

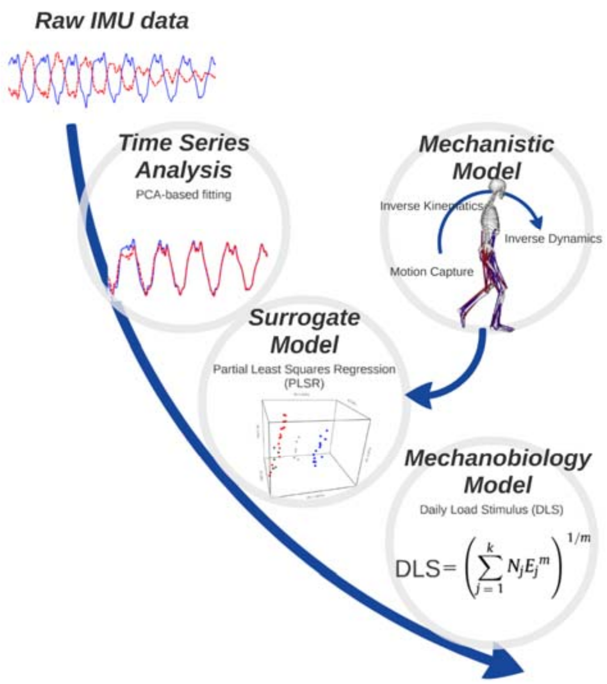

Figure 1: Framework for coupling wearable IMU data with various models to provide estimates of tissue-level loading stimulus.

CONCLUSIONS

We have presented a framework to incorporate wearable sensors with biomechanical models of tissue stress and will present example data collected from basketball players and patients who have had total joint replacement.

ACKNOWLEDGEMENTS

We would like to thank the NZ Medical Technology Centre of Research Excellence for partly supporting this work.

REFERENCES

1. Zhang J et al. *J. Biomech*. **49**:3875-3881, 2016.
2. Beaupré G et al. *J Orthop Res*. **8**:662-70, 1990.

WALKING SPEED, MOTOR CONTROL, AND STRENGTH PREDICT RESPONSE TO SURGERY IN CHILDREN WITH CEREBRAL PALSY

[1]**Apoorva Rajagopal**, [1]Jennifer L. Hicks, [1]Scott L. Delp and [2,3]Michael H. Schwartz
[1]Stanford University, [2]University of Minnesota, [3] Gillette Children's Specialty Healthcare
Corresponding author email: apoorvar@stanford.edu

INTRODUCTION

Cerebral palsy (CP) is a neurologic motor disorder resulting from a brain injury at or near the time of birth. Each injury is unique, and there is large variability between individuals in both the severity and presentation of the movement disorder. Though there are many treatment options including musculoskeletal surgeries to improve the gait of children with CP, identifying good surgical candidates is challenging, and treatment outcomes are variable. Previous studies have shown greater pre-operative gait deformity [1] and better motor control [2] to be predictors of positive treatment outcomes after single-event multi-level surgery (SEMLS). However, because CP patients who undergo a SEMLS are likely to have greater gait deviation than the average CP patient, the data used to train these models represent a skewed subsample of the whole CP population. Thus, it is unknown if the estimated relative effects of various patient characteristics on treatment outcome are biased to better represent the more severely affected patients. In this study, we aimed to (i) build a regression model for predicting patient response to a SEMLS intervention and quantify patient factors that were most predictive of outcome while (ii) correcting for bias in our training data by weighting patients based on their estimated likelihood of having received a SEMLS.

METHODS

We analyzed the affected limb(s) of patients with a diagnosis of CP between the ages of 5 and 18. To be included, the patient had to receive two gait analyses spaced between 9 and 36 months apart with an intervening SEMLS on the affected limb. Normalcy of gait kinematics at the post-surgical gait visit was assessed using the Gait Deviation Index (GDI) [3].

A weighted linear regression model with l_1-regularization to choose a sparse subset of predictive features was computed to predict post-surgical GDI. Candidate predictor variables included features computed from motion capture measurements (gait kinematics, kinetics, and temporal-spatial parameters), physical exam measurements, patient history, and surgical procedures performed. All variables were standardized to have 0-mean and unit variance so that scales of the variables would not affect the l_1-regularization on the regression coefficients. The model was trained using 70% of the observations and tested on the remaining 30%. A 10-fold cross validation using the training data was done to choose the weight for the l_1-regularization on the regression coefficients. To correct for potential sampling bias in the data used to train our regression model, we used inverse propensity score weighting [4]. The propensity scores, or likelihood of receiving a SEMLS based on pre-operative gait variables, were calculated with a random forest classifier trained using the same candidate features as the regression model. Due to the stochastic nature of the model-building process, we averaged effect sizes of features in the linear prediction model over 100 runs.

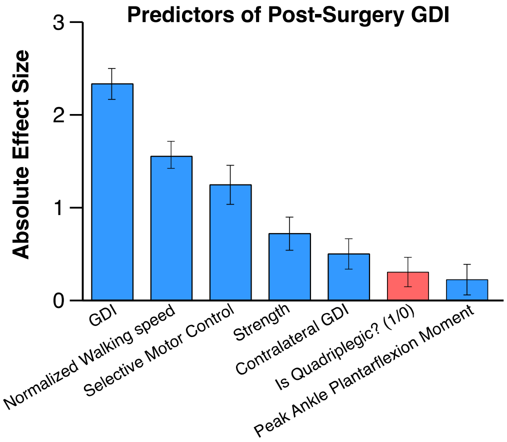

Figure 1: Effect size of pre-operative features in predicting post-operative Gait Deviation Index (GDI). Bars in blue indicate a (+) effect, and bars in red indicate a (-) effect. Height and error bars represent average and standard deviation of effect size, respectively, over multiple models.

RESULTS AND DISCUSSION

We analyzed 1,133 limbs that met all inclusion criteria. The final regression models to predict post-operative GDI included on average 9 variables. Effect size was defined as the change in post-SEMLS GDI per one-point change in the standardized predictor variable. The seven variables with the largest effect sizes were included in over 80% of the regression models (Fig. 1). The models accounted for, on average, 43% of variance in the training data and 40% of the variance in the test data. This performance is comparable to previously published models (e.g., [2]), even with our more conservative modeling approach.

CONCLUSIONS

Propensity score–based weighting provides a method to correct for sampling bias in observational data without sacrificing model performance. While pre-surgical GDI was the strongest predictor of post-surgical GDI, other patient variables including normalized walking speed, selective motor control, and strength were also strong predictors of post-surgical gait kinematics normalcy and were all stronger predictors than any specific surgical procedure. These results suggest that common "patient-intrinsic" factors play a role in determining surgical outcome regardless of the specific surgical procedures performed and should be considered when planning treatment.

ACKNOWLEDGEMENTS

This work was supported by the Mobilize Center (NIH BD2K grant U54EB020405).

REFERENCES

1. Rutz, E, et al. *Gait Posture*. **38**:455-460
2. Schwartz M, et al., *Dev Med Child Neurol.* **58**:1139-1145, 2016
3. Schwartz M, et al. *Gait Posture*. **28**:351-7, 2008.
4. Austin, P, *Multivar Behav Res.* **46**:399-424

ADVANCING A WORLDWIDE RESEARCH NETWORK AND DATABASE TO TRANSFORM BIOMECHANICAL GAIT RESEARCH

[1,2,3] **Reed Ferber**, [1,3] Sean T Osis
[1] Faculty of Kinesiology, University of Calgary
[2] Faculty of Nursing, University of Calgary
[3] Running Injury Clinic
Corresponding author email: rferber@ucalgary.ca

INTRODUCTION

Progress in data science methods, and the ability to collect, store and manipulate "big data," has the potential to improve biomechanical research methods and test broad hypotheses about biomechanical risk factors associated with walking and running gait-related musculoskeletal injury [1]. However, to build accurate classification models, an adequate number of samples are needed, which grows exponentially with the number of features used in the analysis. To directly meet this need, our group has developed the infrastructure and established a worldwide and growing network of clinical and research partners, all linked through an automated three-dimensional (3D) biomechanical gait data collection system: 3D GAIT.

METHODS

The 3D GAIT system is a deployed turnkey motion capture platform specifically designed for gait analysis using a treadmill. The overall system design is a nexus of three main principles: ease-of-use/automation, biomechanics best practices, and data science best practices. Consequently, the system uses off-the-shelf passive motion capture technology, consisting of between three and six infrared cameras (Vicon Motion Systems, Oxford) along with retroreflective markers that are pre-configured for ease of placement on the subject.

The 3D GAIT system derives a "characteristic" pattern from a spatio-temporal normalized set of gait cycles, which are segmented using a machine learning approach to account for inter-subject variability in technique [2]. These normalized gait cycles can then be analyzed by: 1) collapsing into a single representative time-series data set by various averaging techniques (i.e., median, weighted nearest-neighbor interpolant), and/or 2) extracting discrete features from each cycle separately, and merging into a representative feature set for a given subject (i.e., median peak angle, mean angular excursion).

The clinicians and biomechanics researchers who use the 3D GAIT system consists of a growing worldwide network spanning Canada, the USA, Brazil, the UK, Spain, the Netherlands, Australia, New Zealand, and South Korea. All clinical, demographic, and biomechanical data are automatically deposited into a centralized database which can then be applied to hypothesis-driven research.

RESULTS AND DISCUSSION

Despite the advancement of the 3D GAIT system, there are limitations. For example, variability in kinematic variables may be attributed to measurement error, skin marker movement, and inconsistencies in marker placement. Thus, we developed a novel software tool that uses real-time feedback to improve anatomical marker placement [3]. By directly addressing specific limitations, and developing novel solutions using a data science approach, we work towards achieving one of our overarching goals of providing new insights about how to improve gait biomechanics methods.

This mindset can equally be applied to improve clinical practice. For example, our research [4] identified two distinct subgroups of kinematic running gait patterns in healthy runners using a hierarchical cluster analysis without reliance on *a priori* knowledge. The results revealed that one cannot assume that running gait patterns within a sample will be representative of a population - a data science approach was able to reveal this insight.

We have also studied the use of pre-intervention gait kinematics and patient-reported outcome (PRO) measures to predict post-intervention response to a 6-week hip strengthening exercise intervention in patients with knee osteoarthritis (OA) [5]. The result: a unique combination of PRO measures and kinematic variables successfully sub-grouped knee OA patients with a cross-validated classification accuracy of 85.4%.

A final example involves new methods we developed to predict the timing of gait events, using only kinematics, during walking and running, and agnostic of forefoot or heel-toe running gait patterns [2]. This work represents a flexible yet unified approach that is *independent* of specific motion capture technology and can therefore facilitate and standardize the analysis of large volumes of data. We have since integrated the aforementioned techniques into the 3D GAIT software and they are now used across the network.

CONCLUSIONS

The development of our research-to-clinic knowledge translation paradigm is unique within our field and work has just begun to unlock the potential of applying data science methods to help answer complex clinical and biomechanical questions. We encourage the research community to openly share data, employ data science statistical methods, develop similar data research networks, and join our efforts to improve the field of gait biomechanics research.

ACKNOWLEDGEMENTS

This work has been funded by Alberta Innovates Technology Futures, the Canada Research Council, and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Ferber, J Biomech. 49(16): 3759-3761. 2016.
2. Osis, Gait Posture. 46: 86-90. 2016.
3. Osis S, *PLoS One*. 11(1): e0147111, 2016.
4. Phinyomark, J Biomech. 48(14): 3897-3904. 2015.
5. Kobsar D, *PLoS One*. 10(10): e0139923, 2015.