INTRODUCTION

Machine Learning Prediction could Enable Proactive Treatment

- Tools to predict development of disorders resulting from battlefield-related mental or physical trauma are needed to achieve early diagnosis and treatment, with the goal of effective patient outcomes.
- Battelle’s recent work in developing advanced analytics tools to identify, model, and validate early disease prediction biomarkers incorporates machine learning (ML) tools to extract relevant features from health data and predict disease onset and progression.
- This methodology can be leveraged to predict development of warfighter illness and disability, and enable proactive treatment.
- Prototype ML models were developed for knee osteoarthritis (OA):
  - Debilitating disorder caused by degradation of cartilage cushioning the bones of the knee joint
  - Earliest onset can occur during youth, but progresses undetected until the disease has advanced, resulting in knee replacement
- Prototype ML-based tool predicts the probability of knee replacement due to OA.

METHODS

- Data from the Osteoarthritis Initiative (OAI), an online, open source data repository containing data from a multiyear longitudinal study, consisting of 4,796 volunteers
  - Data collected every 12 months for nine years
  - Includes:
    - Image data (MRI and x-ray), 12 TB
    - Non-image data: biomarkers, joint function/symptoms, medical history, nutrition, physical examinations, risk factors
    - Proteomic data subset: Serum and urine protein markers
    - Bone properties and accelerometry

TECHNICAL APPROACH & RESULTS

Clinical and Proteomic Data can Predict Knee Replacement

- Random forest models were used to assess prediction error and find important variables at baseline
- Clinical data: 811 potential covariates, 3,484 subjects
- Biomarkers and Subject Characteristic data - most predictive Composite OA knee grade (L & R) - most important variables
- Proteomic data: 443 potential covariates, 506 subjects
  - MOAKS (MRI OA Knee Score) data - most predictive
  - Variables related to cartilage thickness and morphology - most important
- Subset of 506 subjects: fit random forest models with clinical data, proteomic data, or both:

<table>
<thead>
<tr>
<th>Data</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Only</td>
<td>86.5%</td>
</tr>
<tr>
<td>Proteomic Only</td>
<td>88.1%</td>
</tr>
<tr>
<td>Both</td>
<td>88.5%</td>
</tr>
</tbody>
</table>

Image Data can Predict Knee Replacement

- Convolutional Neural Network (CNN) was trained using patient MRIs:
  - MRI segments from 40 patients to train the CNN
  - Data from 17 patients used to validate the model
  - In total, 509 MRI segments were used for training, and 213 were used to validate (data chunking).

- Using this method, the current best model has 82% prediction accuracy.
- In the validation set, 10 patients did not undergo knee replacement while 7 did, and the model only misclassified 3 of those 7 as not undergoing knee surgery.

CONCLUSIONS

- Battelle’s multi-tiered approach can serve as a new paradigm for the diagnosis and proactive treatment of disease.
- The same methodology being employed for OA can also be leveraged to predict development of warfighter illness and disability due to trauma or exposure.
- Specific treatment based on personalized molecular profiles can then be initiated before symptoms are observed and disease or disability has progressed.