Predictive and Similarity Analytics for Healthcare
Paul Hake, MSPA
IBM Smarter Care Analytics
Disease Progression & Cost of Care

- Healthy / Low Risk
- At Risk
- High Risk
- Early Clinical Symptoms
- Active Disease

20% of people generate 80% of costs
PREDICTIVE MODELING
Problem Definition: Early Detection of Heart Failure (HF)

- **Goal:**
  - How to build a model for predicting HF onset \(x\) months before the HF diagnosis?

- **Data: Longitudinal patient records**
  - Structured data:
    - Demographics, Outpatient diagnoses, Problem List, Vitals, Medication, Labs
  - Unstructured text: encounter notes
What are the known signs and symptoms of HF?

<table>
<thead>
<tr>
<th>Framingham Risk Criteria for Heart Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Criteria</td>
</tr>
<tr>
<td>Paroxysmal nocturnal dyspnea or orthopnea</td>
</tr>
<tr>
<td>Neck vein distention</td>
</tr>
<tr>
<td>Rales</td>
</tr>
<tr>
<td>Radiographic cardiomegaly</td>
</tr>
<tr>
<td>Acute pulmonary edema</td>
</tr>
<tr>
<td>S3 gallop</td>
</tr>
<tr>
<td>Central venous pressure &gt; 16 cm of H₂O</td>
</tr>
<tr>
<td>Circulation time of 25 seconds</td>
</tr>
<tr>
<td>Hepatojugular reflux</td>
</tr>
<tr>
<td>Weight loss of 4.5 kg in 5 days, in response to Rx</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Minor Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilateral ankle edema</td>
</tr>
<tr>
<td>Nocturnal cough</td>
</tr>
<tr>
<td>Dyspnea on ordinary exertion</td>
</tr>
<tr>
<td>Hepatomegaly</td>
</tr>
<tr>
<td>Pleural effusion</td>
</tr>
<tr>
<td>A decrease in vital capacity by 1/3 of max</td>
</tr>
<tr>
<td>Tachycardia (rate of ≥ 120/min)</td>
</tr>
</tbody>
</table>

- Framingham criteria for HF* are common signs and symptoms that are documented even at primary care visits

How predictive are Framingham criteria?

- The prevalence of Framingham criteria varied widely between cases (<1% - 65%) and controls (<1% - 28%)

- The most common Framingham criteria of HF were ankle edema and DOE, but these were also the most common findings in controls, albeit with ~half the prevalence.
Predictive Modeling Pipeline

Structured EHR
→ Feature extraction
→ Universal Feature Model (UFM)
→ Feature construction
→ Feature selection
→ Classification

Unstructured EHR
→ Feature Annotation
→ Universal Feature Model (UFM)
→ Feature construction
→ Feature selection
→ Classification

models

scoring

training
Combining Knowledge and Data Driven Insights for Feature Selection$^1,2$


Method for combining knowledge- and data-driven risk factors

AUC significantly improves as complementary data driven risk factors are added into existing knowledge based risk factors.

A significant AUC increase occurs when we add first 50 data driven features.
9 out of 10 are considered relevant to HF, and one possibly relevant, which confirm the interpretability of the proposed method for expanding knowledge driven risk factors.

The additional features are mostly from medications and symptoms which are complementary to the existing diagnosis (knowledge-driven) features

Table 1: Top 10 data driven features among Cases and Controls

<table>
<thead>
<tr>
<th>Feature type</th>
<th>Feature name</th>
<th>Relevancy to HF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis</td>
<td>DYSLIPIDEMIA</td>
<td>Yes</td>
</tr>
<tr>
<td>Medication</td>
<td>Thiazides and Thiazide-Like Diuretics</td>
<td>Yes</td>
</tr>
<tr>
<td>Medication</td>
<td>Antihypertensive Combinations</td>
<td>Yes</td>
</tr>
<tr>
<td>Medication</td>
<td>Aminopenicillins</td>
<td>Yes</td>
</tr>
<tr>
<td>Medication</td>
<td>Bone Density Regulators</td>
<td>Possible side effect, or maybe a surrogate for elderly women</td>
</tr>
<tr>
<td>Medication</td>
<td>NATRIURETIC PEPTIDE</td>
<td>Yes</td>
</tr>
<tr>
<td>Symptoms</td>
<td>Denial Rales</td>
<td>Yes</td>
</tr>
<tr>
<td>Medication</td>
<td>Diuretic Combinations</td>
<td>Yes</td>
</tr>
<tr>
<td>Symptoms</td>
<td>Denial S3Gallop</td>
<td>Yes</td>
</tr>
<tr>
<td>Medication</td>
<td>Nonsteroidal Anti-inflammatory Agents (NSAIDs)</td>
<td>Yes, contribute to fluid retention due to renal effects</td>
</tr>
</tbody>
</table>
Evaluation Design for Predictive Modeling

- **Diagnosis date**: the day that patient x has been diagnosed with HF
- **Index date**: the day that we want to predict the risk of HF for a given patient x
- **Prediction window**: the time interval between diagnosis date and index date
- **Observation window**: a fixed time interval prior to index date
- **Metric**: Area under the ROC curve (AUC)
Feature-based Patient Representation

- Patients are modeled as longitudinal streams.
- At any time T (indicated by red arrows) for a patient P, we can construct a feature vector to represent the characteristics of P at T.

Remarks
- Absolute time is patient specific. It is not meaningful to compare across patients based on the absolute time.
  - E.g. It does not make sense to compare two patients on their condition at 1/1/2011 in general.
- Relative time is meaningful across patients.
  - E.g. We can compare patients with respect to multiple sequential events, such as a certain medication followed by certain lab results within a month.
- Feature vectors are global. i.e., we can compare and build models on the feature vectors across patients.
Area under the ROC curve (AUC) measure on different prediction windows

- Setting: observation window = 12 months, classifiers = {random forest, logistic regression}, evaluation mechanism = 10-fold cross-validation

- Observation:
  - AUC slowly decreases as the prediction window increases
AUC measure on different observation windows

- Setting: prediction window = 180 days, classifiers = \{random forest, logistic regression\}, evaluation mechanism = 10-fold cross-validation

- Observation:
  - AUC increases as the observation window increases. i.e., more data for a longer period of time will lead to better performance of the predictive model
  - Combined features performed the best at .85 AUC for observation window = 24 months
PATIENT SIMILARITY
Patient Similarity Analytics

**Objective**
Given an index patient, find clinically similar patients for decision support and Comparative Effectiveness

**Highlights**
- Analytics pipeline for similarity that allows flexible combination of information from heterogeneous data sources
- Data driven customization to fine tune similarity metric to specific investigation

Clinical Decision Support

Knowledge Driven
- PubMed
- CPG Clearing House

Data Driven

Knowledge Repository

Average Patient

Personalized
Patient Similarity for Treatment Comparison

Index patient → Patient Indicators → Similarity Metric → Similar patients → Treatment cohort A → Outcome comparison → Treatment cohort B
Analytics Pipeline for Patient Similarity

**Baseline Similarity**
Factors combined using expert defined weights

**Customized Similarity**
Learned context and end point specific distance metric tailored to a specific purpose (outcome, diagnosis, utilization etc.)

Published at: AMIA’10, ICPR’10, ICDM’10, SDM’11a, SDM’11b
**Objective**
Predict the likely outcome of a (patient, physician) pair based on population data and past outcomes

**Highlights**
- Patient and physician characterization using records of past practices and outcomes
- Prediction by analyzing how index patient relates to past success and failure cases of particular physician
- Provides individualized insight vs. population level averages

---

**Physician Outcome Model**

**Comprehensive Diabetes Care (CDC)**
- Born HEDIS 1999
- Percentage of members 18-75 with diabetes (type 1 and 2) who had:
  - HbA1c testing
  - HbA1c poorly controlled (greater than 9%)
  - Retinal eye exam
  - LDL-C screening performed
  - LDL-C controlled (below 130 mg/dL)
  - LDL-C controlled (below 100 mg/dL)
  - Kidney disease (nephropathy) monitored

**Assessment at Population Level**

**Individual Outcome Based**

**Personalized Matching**

**Physician Assessment and Selection**

**Population Based**

Physician Outcome Model

Predict likely outcome based on patient characteristic, provider characteristics and care history

???

patient

???
Problem Formulation

Data
- Diabetic patient’s longitudinal data and their PCPs
- Segmented by patient into baseline condition assessment period and treatment evaluation period
- Used to train and validate models

Samples
- Patients having at least one abnormal HbA1C test result (baseline)

Outcomes
- HbA1C range change between reference and evaluation date (1 year ± 2 months)
  - Positive outcome: range change closer to normal, or remain in “well controlled” range
  - Negative outcome: range change further away from normal, or remain in moderately or sub-optimally controlled

Reference date: one day after the first abnormal HbA1C lab test
Outcome Prediction Process

Physician related features improves prediction for challenging patients

Identifying Challenging Patients

Differentiating Physicians

Total: 195, positive: 81, negative: 114; 80 physicians

Well managed Patients (Positive)

Poorly managed Patients (Negative)

Experiments confirmed that choice of physician has statistically significant impact on challenging patients’ likely outcome
**Objectives**
Continuously assess salient utilization patterns within patient population and how they relate to clinical characteristics; Identify patients with abnormal utilization

**Highlights**
- Identification of dominant utilization groups through patient segmentation
- Specialized predictive modeling methodology linking clinical characteristics to expected utilization
- Identification of unexpected cases via comparison between expected and actual utilization groups for each patient
Utilization Pattern Analysis

Patient population under care in medical home

Patient utilization profiling

Identify patient cohorts with similar utilizations
Unexpected Utilization Detection

Utilization Cohort 1

Utilization Cohort 2

Model 1

Model 2

Model K

Model Aggregation

Predicted Utilization

• Clinical Characteristics
• Demographic Features

Actual Utilization

Unexpected Utilization Detection

Patient Population
Detected Unexpected Utilizations

**Example 1: unexpectedly high utilization**

27 year old female  
Diagnoses:  
HCC127 (Other Ear, Nose, Throat and Mouth Disorders)  
HCC183 (Screening/Observation/Special Exams)

**Example 2: unexpectedly low utilization**

73 year old male  
Diagnoses:  
HCC080 (Congestive Heart Failure)  
HCC166 (Major Symptoms, Abnormalities)  
HCC091 (Hypertension)  
HCC179 (Post-Surgical States/Aftercare/Elective)  
HCC019 (Diabetes with No or Unspecified Complications)  
......

ADVANCED VISUALIZATION
Outflow’s Visual Encoding

- **Past**
- **Now**
- **Future**

- Horizontal position shows sequence of states.
- Color is outcome measure.
- Width is duration of transition.
- Height is number of people.