Session 94PD, Beyond Risk Identification: Predictive Analytics in Health

**Presenters:**
Elena V. Black, FSA, EA, MAAA, FCA
Yi-Ling Lin, FSA, MAAA, FCA
Michael Y. Xiao, FSA, CERA, MAAA

[SOA Antitrust Disclaimer](#)
[SOA Presentation Disclaimer](#)
2018 SOA Health Meeting

YI-LING LIN, FSA, MAAA, FCA

Session 94 – Beyond Risk Identification: Predictive Analytics in Health
June 26, 2018
SOCIETY OF ACTUARIES
Antitrust Compliance Guidelines

Active participation in the Society of Actuaries is an important aspect of membership. While the positive contributions of professional societies and associations are well-recognized and encouraged, association activities are vulnerable to close antitrust scrutiny. By their very nature, associations bring together industry competitors and other market participants.

The United States antitrust laws aim to protect consumers by preserving the free economy and prohibiting anti-competitive business practices; they promote competition. There are both state and federal antitrust laws, although state antitrust laws closely follow federal law. The Sherman Act, is the primary U.S. antitrust law pertaining to association activities. The Sherman Act prohibits every contract, combination or conspiracy that places an unreasonable restraint on trade. There are, however, some activities that are illegal under all circumstances, such as price fixing, market allocation and collusive bidding.

There is no safe harbor under the antitrust law for professional association activities. Therefore, association meeting participants should refrain from discussing any activity that could potentially be construed as having an anti-competitive effect. Discussions relating to product or service pricing, market allocations, membership restrictions, product standardization or other conditions on trade could arguably be perceived as a restraint on trade and may expose the SOA and its members to antitrust enforcement procedures.

While participating in all SOA in person meetings, webinars, teleconferences or side discussions, you should avoid discussing competitively sensitive information with competitors and follow these guidelines:

- Do not discuss prices for services or products or anything else that might affect prices
- Do not discuss what you or other entities plan to do in a particular geographic or product markets or with particular customers.
- Do not speak on behalf of the SOA or any of its committees unless specifically authorized to do so.
- Do leave a meeting where any anticompetitive pricing or market allocation discussion occurs.
- Do alert SOA staff and/or legal counsel to any concerning discussions.
- Do consult with legal counsel before raising any matter or making a statement that may involve competitively sensitive information.

Adherence to these guidelines involves not only avoidance of antitrust violations, but avoidance of behavior which might be so construed. These guidelines only provide an overview of prohibited activities. SOA legal counsel reviews meeting agenda and materials as deemed appropriate and any discussion that departs from the formal agenda should be scrutinized carefully. Antitrust compliance is everyone’s responsibility; however, please seek legal counsel if you have any questions or concerns.
Presentation Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the participants individually and, unless expressly stated to the contrary, are not the opinion or position of the Society of Actuaries, its cosponsors or its committees. The Society of Actuaries does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented. Attendees should note that the sessions are audio-recorded and may be published in various media, including print, audio and video formats without further notice.
Data Analytics in Health

What is currently being done?

- Pricing
- Claims reserving
- Plan design modeling
- Trend forecasting
- Risk scoring
- Care management targeting/savings estimates
- Stress testing
- Data reporting

What can be done?
Applying Data Analytics to Business Problems

Spectrum of data analytics: hindsight to insight to foresight

Adapted from Gartner’s Data Analytics Maturity Model
Types of Problems/Models

- **Linear regression/logistical modeling**
  - Risk adjustment
  - Plan choice modeling
  - Product conversion

- **Survival/Markov models**
  - Disease progression
  - Claims reserving

- **Classification/clustering**
  - Provider referral patterns
  - Targeted marketing
  - Fraud identification
  - High claimant identification

- **Time Series**
  - Trend forecasting
  - Stress testing
An employer group wants to change the medical plans it offers to employees.

**Case Study: Choice Modeling**

<table>
<thead>
<tr>
<th>Current Plan Options</th>
<th>Actuarial Value</th>
<th>Enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMO</td>
<td>0.90</td>
<td>10%</td>
</tr>
<tr>
<td>High Value</td>
<td>0.86</td>
<td>42%</td>
</tr>
<tr>
<td>Medium Value</td>
<td>0.81</td>
<td>45%</td>
</tr>
<tr>
<td>CDHP</td>
<td>0.76</td>
<td>2%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>New Plan Options</th>
<th>Actuarial Value</th>
<th>Enrollment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plan A</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>Plan B</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Plan C</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Plan D</td>
<td>0.68</td>
<td></td>
</tr>
</tbody>
</table>

Case study for illustrative purposes only.
Subject Matter Expertise is Critical

<table>
<thead>
<tr>
<th>ID</th>
<th>Medical Coverage</th>
<th>Coverage Tier</th>
<th>Zip Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Medium Value Option</td>
<td>Employee + 1 Dependent</td>
<td>95066</td>
</tr>
<tr>
<td>6</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>98053</td>
</tr>
<tr>
<td>7</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>60630</td>
</tr>
<tr>
<td>8</td>
<td>HMO Plan</td>
<td>Employee Only</td>
<td>95121</td>
</tr>
<tr>
<td>9</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>33472</td>
</tr>
<tr>
<td>14</td>
<td>HMO Plan</td>
<td>Employee + 1 Dependent</td>
<td>94610</td>
</tr>
<tr>
<td>15</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>94109</td>
</tr>
<tr>
<td>16</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>60478</td>
</tr>
<tr>
<td>17</td>
<td>HMO Plan</td>
<td>Employee + 1 Dependent</td>
<td>94607</td>
</tr>
<tr>
<td>21</td>
<td>Medium Value Option</td>
<td>Employee + 1 Dependent</td>
<td>13777</td>
</tr>
<tr>
<td>24</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>94587</td>
</tr>
<tr>
<td>25</td>
<td>HMO Plan</td>
<td>Employee Only</td>
<td>94109</td>
</tr>
<tr>
<td>26</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>94606</td>
</tr>
<tr>
<td>27</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>94133</td>
</tr>
<tr>
<td>28</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>60532</td>
</tr>
<tr>
<td>31</td>
<td>Medium Value Option</td>
<td>Employee + 2 or More Dependents</td>
<td>95635</td>
</tr>
<tr>
<td>32</td>
<td>HMO Plan</td>
<td>Employee + 1 Dependent</td>
<td>91206</td>
</tr>
<tr>
<td>33</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>02458</td>
</tr>
<tr>
<td>34</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>94103</td>
</tr>
<tr>
<td>36</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>98123</td>
</tr>
<tr>
<td>39</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>60614</td>
</tr>
<tr>
<td>45</td>
<td>Medium Value Option</td>
<td>Employee + 1 Dependent</td>
<td>10604</td>
</tr>
<tr>
<td>48</td>
<td>HMO Plan</td>
<td>Employee Only</td>
<td>94122</td>
</tr>
<tr>
<td>53</td>
<td>HMO Plan</td>
<td>Employee Only</td>
<td>94122</td>
</tr>
<tr>
<td>54</td>
<td>HMO Plan</td>
<td>Employee Only</td>
<td>95356</td>
</tr>
<tr>
<td>55</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>33026</td>
</tr>
<tr>
<td>58</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>94107</td>
</tr>
<tr>
<td>59</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>94123</td>
</tr>
<tr>
<td>61</td>
<td>HMO Plan</td>
<td>Employee Only</td>
<td>94901</td>
</tr>
<tr>
<td>63</td>
<td>HMO Plan</td>
<td>Employee + 2 or More Dependents</td>
<td>94905</td>
</tr>
<tr>
<td>69</td>
<td>Medium Value Option</td>
<td>Employee + 1 Dependent</td>
<td>32703</td>
</tr>
<tr>
<td>72</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>90046</td>
</tr>
<tr>
<td>73</td>
<td>HMO Plan</td>
<td>Employee + 2 or More Dependents</td>
<td>91367</td>
</tr>
<tr>
<td>77</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>94587</td>
</tr>
<tr>
<td>81</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>28173</td>
</tr>
<tr>
<td>82</td>
<td>HMO Plan</td>
<td>Employee Only</td>
<td>95691</td>
</tr>
<tr>
<td>85</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>10573</td>
</tr>
<tr>
<td>89</td>
<td>Medium Value Option</td>
<td>Employee Only</td>
<td>94115</td>
</tr>
</tbody>
</table>

Case study for illustrative purposes only
Importance of Data Visualizations

Case study for illustrative purposes only
Feature Engineering

Case study for illustrative purposes only
Process of Predicting and Evaluating Choice

- Age / Stage in Life
- Risk Tolerance
- Premiums/Contributions ($ and % of Pay)
- Expected Claims
- Plan Design

Individual Plan Election

Individual Annual Total Claims

Total Employee Cost Sharing for the Individual

Plan Cost
Heterogeneous Logit Model

- \( i \) – individuals
- \( j \) - plan options
- \( k \) - # of attributes with weights \( \beta_{ik} \)
- \( U_{ij} \) – utility of plan option \( j \) to person \( i \)
  \[
  U_{ij} = \alpha_i + X_j \beta_i + \varepsilon_{ij}
  \]
  \[
  \beta_{ik} = \beta_{0k} + \beta_{1k} S_{ik} + \sigma_k \mu_k
  \]
- Monte Carlo simulation and maximize log likelihood function
Modeling Approach – New Choices

Know preferences ($\alpha$, $\beta$ and $\sigma$) and now changing the attributes ($X$)

- More Monte Carlo to estimate probabilities
Model Results

Case study for illustrative purposes only
Examining the Range of Results

Case study for illustrative purposes only
Interpreting Results for Business Intelligence

Case study for illustrative purposes only
Now What?

- Evaluate the Model on New Data
- Refine the Model
- Add New Features/Variables
- Prescriptive Analytics
SOCIAL NETWORK ANALYSIS IN HEALTHCARE
What is social network analysis in healthcare and how do we define a relationship?

**Two main relationship types:**
1) Physicians that share patients with other physicians;
2) Physicians that share patients with facilities.
What is social network analysis in healthcare and how do we define a relationship?

**Physician to Physician**

**Physician to Facility**

**Two main relationship types:**
1) Physicians that share patients with other physicians;
2) Physicians that share patients with facilities.
At what level do we define a “shared patient”?

**Patient Level**

**Episode of Care Level**

60% of patients that receive care each year have at least 2 episodes of care per year. There is a significantly clearer relationship of care at the episode level.
Clinically related claims for a single patient are grouped together across a period of time.

SOURCE: Internal Data.
Why is social network analysis for episodes important?

1. Pareto Principle of Healthcare (80/20) Roughly Applies to Episodes
   - Patients with 3 or more episodes are 20% of the population and account for 60% of the cost.

2. Episodes with 2 or more physicians are 30% of the episodes and account for 70% of the cost.

SOURCE: Internal Data.
Episode of Care Example

Episode: 374 – 1.06
Osteoarthritis of the Knee

Physician A: Family Medicine
Physician B: Internal Med-Rheumatology
Physician C: Family Medicine
Physician D: Orthopedic Surgery

R COMMUNITY HOSPITAL
K COMMUNITY HOSPITAL

PHARMACY A
PHARMACY B

Date: 3/14/2015-11/30/2015

Patient X’s Journey Through Episode 374-1.06

A  B  C  D  D


Office visit with family medicine doctor and diagnosed with "pain in joint and multiple site"

Office Visit with internal medicine-rheumatology with some x ray

Office Visit with Family medicine Doctor and diagnosed with "pain in joint"

Office Visit with Orthopedic Surgery

Surgery with Orthopedic and some additional pathology

TRAMADOL HCL retailed in Pharmacy A

TRAMADOL HCL retailed in Pharmacy B

$200  $50,000

SOURCE: Internal Data.
How do we use this information?

- Drive Better Specialist and Facility “Referrals”
- Understand Patient Migration Patterns
- Convince Stakeholders of Value
- Understand Geographic Patterns of Usage
Dallas – Individual Physician to Physician View

Legend

Node size: Total number of patients
Node color: Physician Efficiency

<table>
<thead>
<tr>
<th>Node Color</th>
<th>Efficiency</th>
<th>Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
<td>75% - 100%</td>
</tr>
<tr>
<td></td>
<td>Med-High</td>
<td>50% - 75%</td>
</tr>
<tr>
<td></td>
<td>Med-Low</td>
<td>25% - 50%</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>0% - 25%</td>
</tr>
</tbody>
</table>

SOURCE: Internal data.
Dallas - Physician to Facility View

Legend:
Node size: Total number of patients
Node color: Physician Efficiency

<table>
<thead>
<tr>
<th>High</th>
<th>75% - 100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Med-High</td>
<td>50% - 75%</td>
</tr>
<tr>
<td>Med-Low</td>
<td>25% - 50%</td>
</tr>
<tr>
<td>Low</td>
<td>0% - 25%</td>
</tr>
</tbody>
</table>

SOURCE: Internal data.
How are communities defined?

**Louvain Modularity**

1. Greedy algorithm that maximizes the modularity within communities and minimizes the modularity between communities

2. 

$$Q = \frac{1}{2m} \sum_{vw} \left[ A_{vw} - \frac{k_v k_w}{2m} \right] \delta(c_v, c_w) = \sum_{i=1}^{c} (e_{ii} - a_i^2)$$

3. Small changes can result in very different communities, but the trade-off is acceptable run-time

Dallas – Physician to Physician Efficiency View (Minimum Shared Patient Threshold)

LEGEND

- Node size = Total cost
- Green = efficient physician
- Red = inefficient physician

SOURCE: Internal Data.
Dallas – Physician to Physician View Detail

LEGEND

Node size = Total cost

Green = efficient physician
Red = inefficient physician
Black = insufficient data

SOURCE: Internal Data.
Dallas – Physician to Physician Alternative (Bad) View (No Minimum Threshold)

SOURCE: Internal data.
Houston – Physician to Facility Efficiency
Interactive View

SOURCE: Internal data.
Houston – Physician to Facility Community Interactive View

SOURCE: Internal data.
Open Source Technology Stack

1. Gephi: static visualizations (11, 13, 14, 15)

2. Python [bokeh + networkx]: interactive visualizations (16, 17)
Questions?