Session 25IF, Make Risk Your Friend – Next Generation Claim Prediction

Moderator/Presenter:
Nickolas J. Ortner, FSA, MAAA

Presenters:
Elena V. Black, FSA, EA, MAAA, FCA
Yi-Ling Lin, FSA, MAAA, FCA
Make Risk Your Friend –

Next Generation Claim Prediction

Nick Ortner, FSA, MAAA
Consulting Actuary

nick.ortner@milliman.com

Milliman – Brookfield, WI
(262) 796-3403
Overview / Today’s Agenda

- History / simple world
- Evolution / revolution
- Current models
- Complex / emerging variables
- Samples
- Prediction and execution
- Regulation / challenges
- Transition
Our Simple (Past) World

- Fee for service
- Funding limits?
- Past = future
- Metrics
  - Participation
  - Attendance
  - Clinical
“Business Case” Revolution

- Insurers paying differently
- Slowing funding spigot
- Employers / missions
- Impact on systems
  - Networks
  - Providers (facilities, physicians)
Transformative Models

- Current models
- Concept mainstreaming
- What don’t we know
Harnessing Complex Variables

- Traditional variables
  - Demographic (age, gender, area)
  - Plan design
  - “Clinical” (diagnosis, Rx)

- Emerging variables
  - Variable interactions/combinations
  - Social determinants
  - Community/connections

- Evolving techniques
Sample Projects

- Insurers
  - Wearables
  - Periodic check-ins and changes

- Employers
  - Long-term sustainability
  - Proactive, with requirements
Sample Projects (continued)

- Emergent care risks
  - Medicare: community/interaction

- Opioid addiction risk
  - Commonalities = heightened risk

- Value of changing measures
Prediction ➔ Execution

- Gamification = participation

- Tailored messaging
  ✓ “Meet targets where they are”
Regulation and Challenges

- Transparency
- Privacy
- Other challenges
Nick Ortner, FSA, MAAA
Consulting Actuary – Milliman – Brookfield, WI
nick.ortner@milliman.com  (262) 796-3403
2018 SOA Health Meeting

ELENA BLACK, PHD, CFA, FSA, EA, MAAA, FCA
THE TERRY GROUP
Session 25 – Make Risk Your Friend-Next Generation Claim Prediction
June 25, 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.
Risk Scoring in Health, SOA Studies, and Professionalism Issues
Risk Scoring Modeling in Healthcare

Application of data analytics in healthcare for risk scoring

Traditional to emerging

- Methodologies: from linear regression models to machine learning algorithms
- Data: traditional claims and enrollment data (Rx, Dx, demographic, prior year costs, etc.) to new and emerging data, e.g. socio-economic factors
- Types of risk scoring models: concurrent vs. prospective fitting well into data analytics spectrum

Health risk scores are used for variety of purposes

Many sources of uncertainty
Wealth of Information on Risk Scoring Models

SOA studies related to risk scoring models in healthcare

- Uncertainty in Risk Adjustment (2012)
- Nontraditional Variables in Healthcare Risk Adjustment (2013)
- Accuracy of Claims-Based Risk Scoring Models (2016)
- Risk Scoring in Health Insurance: A Primer (2016)
Potential Issues and Professionalism

Exciting things often come with challenges and potential pitfalls

**Professional guidance (list not exhaustive)**

The Code of Professional Conduct and Actuarial Standards of Practice (ASOPs)

- ASOP 12: Risk Classification
- ASOP 23: Data Quality
- ASOP 25: Credibility Procedures
- ASOP 38: Using Models Outside the Actuary’s Area of Expertise
- ASOP 41: Actuarial Communications
- ASOP 45: The Use of Health Status Based Risk Adjustment Methodologies
- Assumptions Setting ASOPs (27, 35)
- Risk ASOP (51) and Modeling ASOP

**Challenges/issues**

- Messy, often high-dimensional with missing values, data and data quality issues
- Potential bias in data
- Use of proxies
- Non-discrimination, security and confidentiality
- Transparency vs. “black box”
- Spurious correlations: correlation vs. causality
- Interpretability and replicability
- Overfitting and overreliance
- Business purpose appropriateness and applicability

... and... many more
Data Analytics Spectrum and Risk Scoring Modeling in Health
Spectrum of Data Analytics

Descriptive analytics

What happened?

Diagnostic analytics

Why did it happen?

Predictive analytics

What will happen?

Prescriptive analytics

What should I do?

Value towards business solutions

Analytical sophistication

Adapted from Gartner’s Data Analytics Maturity Model
Risk Scoring in Healthcare in Data Analytics Spectrum

What happened?
- Healthcare costs dashboards
- Descriptive statistics
- Data clustering

Why did it happen?
- Healthcare cost trends
- Cost driving features
- Concurrent risk scoring modeling

Prospective risk scoring modeling
- Recalibration off-the-shelf risk scores
- Custom risk scoring models

What will happen?

What should I do?
- Risk stratification and care management
- Choice modeling, simulation and optimization

From hindsight to insight to foresight

Adapted from Gartner’s Data Analytics Maturity Model
Calibration of Risk Scoring Models

Calibration to adjust existing models to specific population

Methodologies

- Full calibration (transparent models, e.g. HHS-HCC)
- Residual calibration to same/similar features (e.g. linear regression on demographic and diagnosis variables)
- Ridge regression residual calibration

Custom risk scoring methodologies

- Custom risk scoring models or risk stratification models
- Residual custom off-the-shelf model recalibration
  - Additional variables/features
  - Different modeling techniques
Full Calibration Example

Calibration of HHS-HCC model to specific population

Demographic Profile: off-the-shelf HCC versus custom calibration

Diabetes ~6.5%

Asthma COPD ~4.5%

Major depressive and bipolar disorders ~3%

Linear regression model based on age/gender and condition bins

Case study for illustration purposes only
Custom Off-the-shelf Model Recalibration

Putting model calibration and ensemble concepts together

Ensemble learning

• Improves predictive analytics results by combining several models (“weak learners”)
  o Bagging (variance decrease)
  o Boosting (bias decrease)
  o Stacking (improves predictions)

Custom Recalibration

• Adjusts to specifics of a given population
• Can use off-the-shelf risk scores as inputs (stacking)
• Potentially reflects additional variables
• Can use different methodologies from original off-the-shelf model
• New spin on residual calibration

SOA 2016 paper briefly explored ensemble idea as analytics question
Emerging Programming Paradigm and Model Evaluation
**New Programming Paradigm: Machine Learning**

- Humans input data & answers
- And how to “learn”... and what does it mean to be wrong...
- Example: clustering algorithm or neural networks or decision tree/Random Forest

---

**Traditional Modeling versus Machine Learning**

Could computer automatically learn the rules by looking at data?

**Classical programming model**
- Humans input data and set of rules/function on how to arrive at answers
- Also how close they want data to fit to the “model”...
- Example: linear regression or generalized linear regression

**New Programming Paradigm: Machine Learning**
- Humans input data & answers
- And how to “learn”... and what does it mean to be wrong...
- Example: clustering algorithm or neural networks or decision tree/Random Forest
Examples of ensemble models

- Based on decision trees
  - Random forest: multitude of trees trained (random subsets of data) and results averaged
  - Gradient boosting: trees are trained in succession on residuals of target versus sum of previously trained trees

Decision Trees and Ensemble Methods

Ensemble approaches often result in robust models

Rules based decision tree
- Perfect for classification problems, but can be used for regression
- Transparent and easy to interpret
- Training is done by optimizing given “loss” function
- …. But a “weak” learner

Case study for illustration purposes only
Risk Scoring Model Evaluation

Model evaluation is an important part of any modeling project

- Relevance and importance of criteria
- Appropriate and consistent with purpose
- On “unseen” or “test” sample of data
- Examples of criteria/metrics
  - Standard statistical measures (R squared, RMSE, MAE, etc.)
  - Predictive Ratios: grouped A/E type measures (demographic groups, diagnostic groups, cost groups, random groups, etc.)
  - Tolerance curves
  - ROC curves for Cost Groups
  - Correlation and comparison with naïve and standard models

Cautionary tale!
Famous Anscombe’s quartet: all four datasets have the same statistical properties, including R squared=0.67, means and variance of x and y, correlation and linear regression model: y=3+0.5x
Case Study: Custom Risk Scoring Modeling
Case Study: Descriptive Analytics

Dashboards, distributions, descriptive statistics

This is traditional analysis to inform what actually happened and the first step in any modeling project.

Demographic profile of the population

In this case study babies under age of 2 were excluded, and population shown were enrolled at least for one month in both years.

Case study for illustration purposes only

Average PMPM (Medical and Rx) by year and gender

Case study for illustration purposes only

Shaded area illustrates male percentile

52% Male

42% Male

0% 10% 20% 30% 40% 50% 60%
baby child 18-24 25-29 30-34 35-39 40-44 45-49 50-54 55-59 60-64 65+
4% 12% 6% 5% 7% 7% 8% 9% 10% 9% 5%
Case Study: Diagnostic Analytics

Investigating and identifying trends & relationship

Diagnostics focused on uncovering patterns, relationships, trends, and potentially engineering predictive features

Relationship between potential predictors (independent variable), relationships between predictors and target, potential transformed variables relationships

Claim costs are lognormally distributed: fitting normal distribution to log of PMPM costs for current and prior years → visually there is a linear relationship but correlation is only 0.54
Case Study: Start Simple!

Simple approach: variation on “stacking” concept

Comparison of age-group predictive ratios: Off-the-shelf versus linear regression with HCC diagnosis severity as input

Linear regression on three variable
Predictive Ratios on Test Data by Age Group

Predictive Ratios by Age Group (Test Data)

Case study for illustration purposes only

Off-the-shelf HCC (test data):
$R^2$ is 0.24, and correlation of predicted values versus target is 0.65

Linear regression on three variables (test data):
$R^2$ is 0.42, and correlation of predicted values versus target is 0.65
Case Study: Complexity versus Interpretability

Gradient Boosted Trees or Random Forest: More Accurate-Hard to Interpret

Many machine learning models are hard to explain/interpret

Two models (linear and “bagged trees”) fit to the same variables, but the scatter shown against just one predictor ($R^2 = 0.2$ for linear and 0.3 for random forest)

Feature importance allows for easier interpretation but also predictive power analysis

Case study for illustration purposes only
Case Study: Predictive Analytics

Calibrating residual using bagged trees and additional features

On Test Data:

- $R^2$ is 0.24, correlation 0.65, MAE=67% for off-the-shelf HCC
- $R^2$ is 0.48, correlation 0.70, MAE = 73% for residual custom-recalibrated HCC

Case study for illustration purposes only
Various uses of risk scoring in health care: population health and care management

Identifying best cases for care management:

- **High cost**
  - Acute illness, trauma, accidents
  - Chronic deceases, high cost and risk

- **Low cost**
  - Healthy
  - Rising cost?

- **Risk**
  - From low to high

- **Cost**

Prevention and wellness programs

Case study for illustration purposes only

### Case Study: Decision-informing Analytics

**Quadrant Dashboard**
(Members with Chronic Condition)

<table>
<thead>
<tr>
<th>Quadrant</th>
<th>Low Risk - Low Cost</th>
<th>Low Risk - High Cost</th>
<th>High Risk - Low Cost</th>
<th>High Risk - High Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Score (Average)</td>
<td>0.6869</td>
<td>0.9119</td>
<td>3.0892</td>
<td>6.9507</td>
</tr>
<tr>
<td>PMPM (Average)</td>
<td>$274</td>
<td>$8,956</td>
<td>$708</td>
<td>$5,300</td>
</tr>
<tr>
<td>% of Member</td>
<td>56.27%</td>
<td>1.63%</td>
<td>28.91%</td>
<td>13.19%</td>
</tr>
<tr>
<td>Aggregate Claims (Sum)</td>
<td>$2.1M</td>
<td>$0.6M</td>
<td>$2.7M</td>
<td>$8.3M</td>
</tr>
</tbody>
</table>

Chronic Conditions: Arthritis, Asthma, Cardio, COPD, Depression, Diabetes, High Blood Pressure, Kidney, Obesity, Opioid

Assess characteristics of high risk/low cost group: potential for care management.
Questions? Thoughts... Comments?