SOCIETY OF ACTUARIES

Health Spring Meeting
June 2009

Session # 23 PD: Predictive Modeling Hybrids Gaining Increased Mileage

Robert Bachler, FSA, MAAA, FCAS
Howard Brill, Ph.D.
Ian G. Duncan, FSA, FIA, FCIA, MAAA

Moderator:
Vincent Michael Kane, FSA, MAAA

Predictive Modeling Hybrids

Underwriting High Risk Individuals in Stop Loss

Presented by
Robert Bachler

June 8, 2009

Milliman

Underwriting High Risk Individuals

Overview

- Current Practices
- Hybrid Predictive Model – Concept
- Hybrid Predictive Model - Applications

Milliman

Underwriting High Risk Individuals

Current Practice

- Complexity of reviews vary
  - Underwriter review of key condition(s)
  - Nurse/physician review of medical records
- Pricing approaches vary
  - Pooling
  - Covering Exclusions/Reductions
    - Lasering
    - Aggregating Specific
  - Additional Premium

Underwriting High Risk Individuals

Current Practice - Problems

- Volume of cases vs. cost of review
- Variability of actual future costs
  - Ground-up costs
  - Excess costs
- Unique nature of each individual
  - No one has seen enough cases “like this one”

Underwriting High Risk Individuals

Current Practice - Example

And the clinician (or the underwriter) says…… $150,000 to $200,000

Joe – 54 year old male, diabetic w/CHF, $125K in claims last year

Milliman

Underwriting High Risk Individuals

Current Practice - Example

- If Joe's group has a stop loss deductible of $100,000
  - Underwriter will assume an extra cost of between $50,000 and $100,000
  - Underwriter may modify Joe's deductible to $200,000 laser

Both understate the cost of variability

- If Joe's group has a stop loss deductible of $250,000
  - Many underwriters wouldn't adjust for Joe

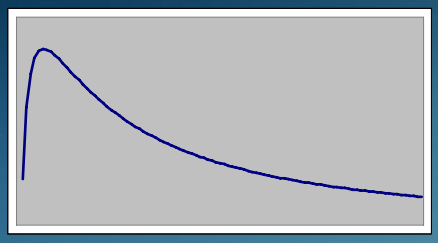
Implicit assumption – Joe has the same probability of a $250,000 claim as an "average" 54 year old male

Milliman

Underwriting High Risk Individuals

A Hybrid Predictive Model

- A new approach



Joe – 54 year old male, diabetic w/CHF, $125K in claims last year

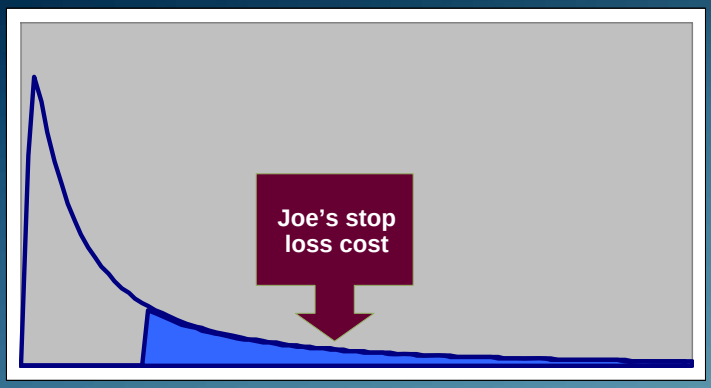
Joe's claim cost is a stochastic variable following a unique distribution

Milliman

Underwriting High Risk Individuals

A Hybrid Predictive Model

- If Joe's group has a stop loss deductible of $100,000



Milliman

Underwriting High Risk Individuals

A Hybrid Predictive Model

- Underwriter's Options for Joe

Joe – 54 year old male, diabetic w/CHF, $125K in claims last year

- Charge premium to cover Joe's higher stop loss cost
- Set a "laser" for Joe so that his expected stop loss cost equals the group rate
- Include Joe's distribution in an "aggregating specific" calculation
- No change, but put Joe's excess cost into a "pool" to cover underwriting decisions

Underwriting High Risk Individuals

Application #1 – Munich Re America and DxCG

- Characteristics of the DxCG model fundamental to application
  - Additive regression coefficients generated using binary independent variables
  - Certain groups of condition categories are assigned hierarchies
  - Methodology open to the addition of other independent variables, such as prior year cost.

Underwriting High Risk Individuals

Application #1 - Munich Re America and DxCG

- Munich Re America application
  - Identify condition categories that are key to generation of stop loss claims
  - Require user to identify which of these conditions are present in the high risk individual of interest
  - Using methods similar to "base" DxCG model, generate an underlying expected cost and estimated variance around that expected cost

Many two-parameter distributions are readily defined by their mean and variance

  - Adjust results for ancillary information (geography, network reimbursement, clinician estimates).

Milliman

Underwriting High Risk Individuals

Application #2 - Milliman

- Characteristics of the "ground up" Milliman model
  - Heavily dependent on binary independent variables like DxCG
  - For select conditions, includes "modifiers" that estimate likelihood of developing condition next year and resulting expected cost
    - Modifiers utilize additional information relevant to the condition, such as precursor conditions and procedures performed in prior year.
  - Also includes ancillary information such as prior year inpatient, outpatient, Rx claims paid

Milliman

Underwriting High Risk Individuals

Application #2 - Milliman

- Milliman application
  - Identify conditions that are key to generation of stop loss claims
  - Require user to identify which of these conditions are present in the high risk individual of interest
  - For key conditions with "modifiers", require user to enter key information
  - Generate an underlying expected cost and estimated variance around that expected cost
  - Adjust results for ancillary information (geography, network reimbursement, clinician estimates).

Same problem, different models, both have a solution.

Underwriting High Risk Individuals

Other models

- Munich Re America and Milliman situations had similar characteristics
  - Methodology transparent to hybrid builder
  - Fairly simple statistical method (linear regression) serves as foundation of underlying models
- What if models are "black boxes"?
  - Use what you have
  - If you can generate risk scores/expected costs, you can bucket individuals and generate empirical distributions.

THINK OUTSIDE THE (BLACK) BOX!!!

Hybrid Predictive Models

Howard Brill, Ph. D.
Monroe Plan for Medical Care
SOA Conference, Toronto, June 8, 2009

MPFMC and Predictive Models

- The Monroe Plan for Medical Care is a Medicaid managed care organization in upstate New York, covering 106,000 lives.
- The Monroe uses both pure commercial off-the-shelf (COTS) models and mixed models combining COTS and customized components.

Hybrid Models

- Definition: A mix of COTS and customized components.
  - The customized components may be created by the vendor, customer, or a third-party.

Monroe Plan Examples

- DxCG DCGs and the Milliman MedInsight BI package.
- DCGs + Diabetes Gaps in Care + Casetrakker Case Management System
- Internally calibrated Likelihood of Hospitalization (LOH) Model.
- Internally calibrated and customized LOH model adapted to predicting ED admissions.
- Internally calibrated and customized model for predicting inpatient readmissions built on DCG HCCs.

Why?

- Unique types of predictors (ethnicity, county, benefit, special clinical measures) or outcomes (readmission)
- Special characteristics of data (Medicaid benefit structures can vary from state to state)
- Maximizing cost-benefit of interventions
- Operational process and integration

Where?

- The predictive model is usually a small part of the overall system.
- With today's industry only some components are meaningful targets for customiztion.

Structure of Modern Predictive Models

- Predictive Models in healthcare have three major components:
  - A *diagnostic grouper*, which is a variable reduction mechanism.
  - A *statistical model*, which estimates the probability or expected value of an outcome of interest.
  - An *operational process* for using the results of the predictive model for intervening in the healthcare system.

What?

- Diagnostic groupers are expensive to build and maintain. They required highly specialized skill sets (statisticians, physicians and code experts) and very large data resources.
- Few plans will have the capability to develop its own groupers and even maintaining published groupers is daunting.

What?

- The effective integration of the predictive model into an operational process will involve customization in all but the smallest plans.

What?

- Calibrating or customizing the statistical model is the key decision point.
  - Comparability: Vendor's calibration data versus target population?
  - Correspondence of vendor's outcome variable to business target
  - Special conditions --- benefits, carve-outs, patient subsets, medical specialities

When?

- Adoption cycle
  - Creating effective operational processes will offer a greater benefit early in the adoption cycle than optimizing the statistical model.
  - Data quality problems and operational dysfunction will have a bigger impact on success than the difference between an off-the-shelf and customized calibration.

How?

- Vendors, consultants, in-house
- Avoid an episode of “biostatisticians gone wild”
  - Focus on conceptualizing the business problem and effectively integrating solutions with underwriters and clinical staff

Criteria for Evaluating Models and Systems

- Statistical
  - Train / Validate Method
  - Explained Variance
    - R^2
    - Outlier Sensitivity
    - Tail Performance
  - Discrimination
    - Classification Error / Specificity and Sensitivity
    - ROC
    - c-index
- Information Criteria: AIC, BIC

Criteria for Evaluating Models and Systems

- Technical and Business
  - Transparency
    - Ease of Explanation
    - Clinical credibility
    - Regulatory credibility
    - Documentation
  - Reproducibility / Stability
    - Impact of model updates
    - Application of model to new populations
  - Robustness
    - Ability to accommodate data quality problems
  - Business Process Flexibility

Summary

- Early in the adoption of predictive models, the off-the-shelf model will be substantially stronger than previous methods.
  - Data quality and process problems will exceed the potential benefit of customization.
- With mature operational processes, recalibration and customization can be a source of innovation.

SCIOinspire
FRESH thinking REAL results

Innovating Healthcare Business Process Service Delivery

Predictive Modeling in Wellness Programs

June, 2009

Agenda

1. Background and Issues.
2. Model Objectives.
3. Population/Data.
4. Sample model.
5. Program Planning.
6. Program Evaluation.
7. Discussion.

Solucia Consulting, *Corporate Information*

- Actuarial Consulting firm founded in 1998. A leader in managed care, disease management and predictive modelling applications.
- Acquired by SCIOinspire Corporation in April 2008.
- 6 healthcare actuaries; 2 PhDs; healthcare analytics team (6).
- Four main businesses:
  - Disease and Care Management consulting (operations; ROI; outcomes; predictive modelling);
  - Actuarial consulting (all actuarial functions for the Freelancers Insurance Co. (NY); state Medicaid plans, etc.); Massachusetts Healthcare Connector Board (member);
  - Software applications.
  - Care management operations support services (Analytics, data management, risk assessment, outreach, fulfilment)
- Strong research and public policy foundation: we have always supported a strong research function to inform our recommendations.

Notes

- This presentation is drawn from Chapter 15 of my forthcoming book: "Managing and Evaluating Healthcare Intervention Programs" (published by Actex, October 2008).
- Authors of this chapter are Ian Duncan, FSA MAAA, Manjula Singh, PhD and Christian Birkmeyer, MS.

Background / Issue

- As wellness programs grow in number, sponsors need methods for evaluating and comparing financial performance between programs.
- Within the industry we generally encounter two different methodologies for evaluating population programs such as wellness.
  - An application of the adjusted historical control methodology (the industry standard for evaluating chronic management programs); and
  - The "risk factor reduction" method.

Background / Issue

1. Adjusted Historical Control Methodology

Advantages	Disadvantages
1. Well-researched and well-accepted in applications such as chronic condition population measurement.	1. Requires an independent trend estimate. Unlike chronic condition evaluations, there is no "natural" population from which to derive a trend estimate.
2. Can be applied to an entire population; does not require additional information (such as self-reported or biometric risk data).	2. Difficult to ensure comparability between populations over time.

2. Risk Factor Reduction Methodology

Advantages	Disadvantages
1. Financial results are directly derived from risk factor reduction in the managed population. Savings may be tied directly to the intervention.	1. Requires a credible volume of data collected at (minimum) two points in time, usually 12 months apart.
	2. Requires financial values appropriate to the observed risk-factor reduction.
	3. May require some element of trend estimation to allow for "natural" progress of risk factors over time.
	4. May suffer from "self-reporting bias" depending on which participants respond, and how.

This presentation discusses work we performed to calculate financial value of risk factors.

Chapter 15

Chapter 15 is in three parts.

In Part 1, we develop a predictive model that relates health risk factors and claims costs.

In Part 2, we illustrate how the model may be used to plan intervention programs and prioritize member interventions.

In Part 3 we illustrate how the relationship between risk factors and claims costs developed in Part 1 may be used to estimate the savings from risk factor reduction due to an intervention program.

Population and Data

- We developed multivariate regression models that predict member claims costs based on independent variables (risk factors).
- The study population consists of members of a commercial health plan (HMO and PPO coverage), with 24 months of continuous enrollment from April 2005 through March 2007, who had completed self-reported Health Risk Assessment (HRA) surveys between May 2005 and March 2006. Self reported health care attributes were available for 18,610 members and includes responses to more than 400 questions on lifestyle factors, family history, and conditions which are not identifiable through claims (e.g. obesity, smoking and other behavioral factors).

Population and Data

The HRA covers the following factors:

- Personal Disease History: presence of a diagnosis of Asthma, Cancer, Depression, Heart Disease, High Blood Pressure, Stroke and other selected high cost diseases.
- Family Disease History: Diagnosis of Cancer, Heart Disease, High Blood Pressure, Stroke etc.
- Health Screenings and Immunizations: Influenza, Pneumonia,...
- Alcohol Consumption: Ability to limit drink in various stressful situations.
- Injury Prevention Behavior: Gun safety, wearing seat belts etc.
- Nutrition: Consumption of grains, nuts, dairy, and portions thereof.
- Physical Activity: low, medium, high intensity.
- Skin Protection: use in outdoor activities.
- Stress and Well-Being: ability to handle stressful situations.
- Tobacco use: cigar vs. pipe vs. cigarette, how many, how addicted.
- Weight Management: Body Mass Index; and
- Women's Health: pregnancy status, receiving hormone replacement therapy.

Population and Data

Mean Claim Costs (April 2006-March 2007) by Gender			
	All	Males	Females
# Members	18,610	8,576	10,034
Mean Cost	$ 3,586	$ 2,974	$ 4,109

Three versions of models were developed with different explanatory variables:

1. DxCG prospective Med+Rx risk only;
2. DxCG prospective Med+Rx risk and HRA variables;
3. HRA variables only.

Model – example (Male)

Male Cost Model Based on Prospective Risk and HRA Variables						
Attribute	Variable	Coefficient		Range of variable	Mean Variable	Contribution to predicted cost at mean value
	Intercept	4,297	0.001	1	1.00	4,297
Risk	DxCG prospective Med+Rx risk	1,865	<.0001	0.16-7	1.36	2,529
Personal Disease History 1	Congestive Heart Failure (CHF), Osteoporosis, Angina, TIA (mini-stroke lasting less than 24 hrs)	9,231	<.0001	0 (No), 1 (Yes)	0.02	157
Weight Management	BMI group	374	<.0001	1 (< 25), 2 (25-29.99), 3 (30-34.99), 4 (35-39.99), 7 (> 40, no value)	2.29	856
Stress and Well-Being	Please rate how confident you are that you can practice stress management techniques consistently.	(553)	0.0065	1-3. No value=3	2.66	(1,473)
Stress and Well-Being	In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?	285	0.0157	1 (Never), 2 (Almost Never), 3 (Sometimes, No value), 4 (Fairly Often), 5 (Very Often)	2.12	604
Tobacco	Please rate how confident you are that you can keep from smoking cigarettes when you feel bored.	(447)	0.0176	1-5, No value=4.5	4.34	(1,937)
Physical Activity	Low-intensity physical activity – days per week	(293)	0.0245	5-7, No value=5	5.70	(1,668)
Nutrition	On a typical day, how many servings of fruit do you eat?	(373)	0.0353	1 (None, no value) 2 (one) 3 (two plus)	2.12	(790)
Family Disease History	Has anyone in your immediate family (father, mother, brother, sister) been diagnosed with Colorectal Cancer	1,134	0.0433	0 (No) 1 (Yes)	0.05	55
Health Screenings	Have you had a Clinical Skin Exam by a health care provider within the last 12 months?	545	0.0466	0 (No) 1 (Yes)	0.26	144

Use in program planning

- Program planning requires understanding of program costs and potential benefits.
- Long-term benefits of wellness programs are not well understood. As a proxy, we use the cost correlations from the model.
- Modifiable variables in the female model include BMI, nutrition, smoking and stress management.
- The costs associated with each variable, in conjunction with the cost of a program, allows the program sponsor to plan optimally for interventions.

Economics of program planning

Example: weight management:

- Opportunity: Δ 1 category = $374.
  - 3,000 eligible members (ee/dep);
  - Percentage overweight: 10%;
  - Percentage enrolling: 20%;
  - Percentage reducing by 1 category: 25%;
  - 10% * 20% * 25% = 0.5% of 3,000 = 15;
  - Potential Dollar Savings: 15 * $374 = $5,610

  → OPPORTUNITY: $5,610.

- Program Cost: Must be < $5,610, or < $1.87 PMPM.

Use in program evaluation

Savings Estimation Based on Female HRA only model						
Attribute	Variable	Cost Coefficient (Table 8)	Mean Variable (Baseline) (Table 8)	Mean Variable (Post Intervention)	Cost (Baseline)	Cost Improvement
(a)	(b)	(c)	(d)	(e)	(f)	(g)
Weight Management	Body Mass Index	$118	29.61	28.00	$349,521	$(18,981)
Physical Activity	Moderate-intensity physical activity - minutes per day	(46)	13.49	15.00	(61,736)	(6,905)
Stress and Well-Being	In the last month, how often have you been angered?	1,632	0.05	0.04	8,700	(2,196)
Physical Activity	High intensity activities? (hours per week)	(306)	0.34	0.50	(10,450)	(4,835)
Nutrition	Servings of grain per day?	(868)	0.99	1.00	(85,937)	(1,220)
Tobacco	Rate confidence to avoid smoking when blue	(294)	1.74	1.74	(51,089)	(120)
All other variables (in table 8)					263,812	-
TOTAL					$412,821	$(34,256)

Discussion