



Joint Colloquium of the IACA, PBSS and IAAHS Sections of the International Actuarial Association

Westin Copley Place Hotel, Boston, U.S.A. – 4-7 May 2008

SOA Sponsored Research Project:

A Comparative Analysis of Claims-Based Tools for Health Risk Assessment

Jonathan Shreve, FSA, MAAA



Agenda

- **SOA Study:**
 - Introduction / Background
 - Results Overview
 - Methodology
 - Detailed Results
- **Question & Answer Period**
- **Further Advancements in Risk Selection – LBA**
 - Uses and relevance of LBA
 - Incorporation of LBA in underwriting and pricing
 - Development of LBA models

Milliman-SOA 2006 Comparative Study of Risk Adjustment Tools

- Purpose to provide unbiased information on the predictive accuracy and other characteristics of commercially available health claims-based risk adjustment tools
- Prior studies done in 1996 & 2002
- Models tested – 7 in 2002, 12 in 2006

Introduction: Health Claims-Based Risk Adjustment

- Claims-Based Risk Adjusters measure morbidity
- Claims data > Disease groupings > Risk weights > Prediction
- Demographic and Pharmacy information is added as independent variables
- Disease groupings rely on diagnosis, utilization, cost and pharmacy information

Introduction: Applications of Risk Adjustment

- Underwriting (Renewal)
- Provider or plan payment
- Provider efficiency & profiling
- Disease management

Evolution of Risk Adjuster Tools

- Multiple models within the same tools to address specific modeling scenarios (such as data lags or components of cost)
- Standard approach:
 1. Assign disease indicators – 1s and 0s in the assessment period
 2. Linearly regress projection period costs onto the disease indicators
- Techniques have become more sophisticated, including non-linear regression and AI techniques
- Addition of prior cost to underwriting models

Vendors / Products Included in Study

- Adjusted Clinical Groups (ACG, ver 7.1)
- Adjusted Clinical Groups – with prior pharmacy cost (ACG)
- Chronic Illness & Disability Payment System (CDPS, ver 2.5)
- Clinical Risk Grouping (CRG, ver. 1.4)
- Diagnostic Cost Groups (DCGs, ver. 2.1.1)
- Episode Risk Groups (ERGs, ver. 5.3)
- Impact Pro (Ingenix)
- MEDai
- MedicaidRx
- Pharmacy Risk Groups (Ingenix, ver. 5.3)
- RxGroups, ver. 2.1.1 (DxCG)
- Underwriting Model, RiskSmart, (DxCG, ver. 2.1.1)

Scenarios / Conditions

- R-squared / MAPE / Predictive Ratios
- Truncation Level (None, 250k, 100k)
- Prospective / Concurrent
- Offered / Recalibrated
- **Non-Lagged / Lagged**
- **With / Without Prior Cost**

Comparison Measures

- **R-Square**

$$\text{R - Square} = 1 - \frac{\sum (\text{Actual} - \text{Predicted})^2}{\sum (\text{Actual} - \text{Average of Actual})^2},$$

- **Mean Absolute Prediction Error**

$$\text{MAPE} = \frac{\sum |\text{Actual} - \text{Predicted}|}{N}.$$

- **Prediction Ratio**

$$\text{Prediction Ratio} = \frac{\text{Sum of Predicted}}{\text{Sum of Actual}}.$$

Individual Results Overview

**Table I.1 - R-Squared and MAPE for Prospective Nonlagged - Offered vs. Optimized
(Recalibrated, with Prior Cost, 250k Claim Truncation)**

				Offered Models		Optimized Models w/ Prior Costs	
Risk Adjuster Tool	Developer	Inputs	R-2	MAPE %	R-2	MAPE %	
ACG	Johns Hopkins	Diag	19.2%	89.9%	23.0%	86.2%	
CDPS	Kronick / UCSD	Diag	14.9%	95.3%	24.6%	85.6%	
Clinical Risk Groups	3M	Diag	17.5%	90.9%	20.5%	86.6%	
DxCG DCG	DxCG	Diag	20.6%	87.5%	26.5%	82.5%	
DxCG RxGroups	DxCG	Rx	20.4%	85.3%	27.1%	80.7%	
Ingenix PRG	Ingenix	Rx	20.5%	85.8%	27.4%	80.9%	
MedicaidRx	Gilmer / UCSD	Rx	15.8%	89.6%	26.3%	81.9%	
Impact Pro	Ingenix	Med+Rx+Use	24.4%	81.8%	27.2%	80.6%	
Ingenix ERG	Ingenix	Med+Rx	19.7%	86.4%	26.5%	81.2%	
ACG - w/ Prior Cost	Johns Hopkins	Diag+\$Rx	22.4%	85.6%	25.4%	82.1%	
DxCG UW Model	DxCG	Diag+\$Total	27.4%	80.4%	29.1%	78.3%	
Service Vendor		Inputs	R-2	MAPE	R-2	MAPE	
MEDai	MEDai	All	N/A	N/A	32.1%	75.2%	

Comparison to Prior Study

Table IV.3 – Comparison to 2002 Study of Offered Weight R-Squared Prospective Nonlagged by Claims Truncation Level

Risk Adjuster Tool	2002 Study		Current Study		Average Change in R2
	100K	None	100K	None	
CDPS	12.5%	10.3%	17.6%	12.4%	2.1%
DCG	18.0%	14.3%	22.3%	17.4%	3.1%
Medicaid Rx	9.8%	7.1%	19.3%	12.9%	5.8%
RxGroups	18.1%	13.4%	23.8%	16.8%	3.4%
ERG	19.3%	14.6%	23.7%	16.2%	1.6%
Average	15.5%	11.9%	21.4%	15.1%	3.2%

Executive Summary of Grouped Results

Disease Based Groupings

**Table I.2 – Predictive Ratios by Medical Condition in 2003
(Offered Nonlagged Prospective, 250K Truncation)**

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	88.4%	100.0%	96.7%	103.1%	99.6%	92.3%
CDPS	Diag	95.0%	73.4%	84.8%	76.4%	67.3%	92.5%
Clinical Risk Groups	Diag	85.1%	94.7%	99.7%	99.5%	91.5%	89.0%
DxCG DCG	Diag	93.3%	98.3%	98.6%	103.2%	86.4%	95.9%
DxCG RxGroups	Rx	95.5%	76.9%	97.9%	89.4%	89.2%	88.6%
Ingenix PRG	Rx	94.9%	93.9%	98.2%	89.7%	79.6%	87.1%
MedicaidRx	Rx	90.1%	94.9%	92.7%	79.1%	90.8%	94.0%
Impact Pro	Med+Rx+Use	97.6%	115.4%	96.4%	99.8%	95.1%	98.0%
Ingenix ERG	Med+Rx	90.0%	99.2%	94.8%	92.9%	80.0%	91.9%
ACG - w/ Prior Cost	Diag+\$Rx	92.5%	109.0%	95.8%	97.5%	103.6%	91.0%
DxCG UW Model	Diag+\$Total	93.2%	84.9%	91.1%	90.7%	103.6%	94.6%
Average		92.3%	94.6%	95.2%	92.9%	89.7%	92.3%

Executive Summary of Grouped Results

Cost-Percentiles Based Groupings

**Table I.3 Prospective Optimized (Recalibrated, with Prior Costs), Nonlagged
Predictive Ratios by Cost Percentile Groupings
(Cost Groupings Defined for 2004)**

Risk Adjuster Tool	Percentile Ranges							
	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	27.1%	46.7%	69.6%	99.1%	146.5%	249.9%	544.2%	8433.1%
CDPS	24.2%	43.8%	67.8%	98.6%	150.4%	256.7%	546.1%	8537.4%
Clinical Risk Groups	28.4%	49.2%	73.0%	103.5%	150.4%	238.8%	488.7%	6808.8%
DxCG DCG	25.2%	45.6%	70.4%	101.1%	149.7%	248.5%	528.7%	7780.7%
DxCG RxGroups	24.9%	48.0%	75.0%	105.4%	151.3%	237.3%	482.6%	7177.5%
Ingenix PRG	25.0%	48.0%	74.5%	104.4%	150.6%	238.0%	489.1%	7426.9%
MedicaidRx	24.2%	46.4%	73.4%	106.2%	155.8%	243.8%	478.5%	6773.7%
Impact Pro	29.7%	50.6%	74.9%	103.6%	149.5%	235.0%	470.1%	6587.2%
Ingenix ERG	24.3%	46.1%	73.6%	107.4%	156.4%	245.1%	482.0%	6226.3%
ACG - w/ Prior Cost	27.2%	51.7%	76.5%	102.1%	141.7%	230.3%	510.3%	8146.4%
DxCG UW Model	26.8%	50.9%	77.4%	107.6%	150.4%	229.0%	452.4%	6427.8%
MEDai	29.5%	52.5%	78.0%	106.5%	145.4%	216.2%	411.9%	5592.5%
Average	26.4%	48.3%	73.7%	103.8%	149.8%	239.0%	490.4%	7159.9%
Age-Gender Only	7.0%	20.0%	42.0%	77.0%	153.0%	327.0%	861.0%	17443.0%

Overview of Methodology

- Data Used
- 50/50 design
- Calibration
- Comparison measures (R-Square / MAPE / Predictive Ratios)

Data Used

- MarketScan® MedStat Database
- About 620,000 members with \$3b in incurred claims
- Non-Lagged Analysis: experience period is 2003, projection year is 2004
- Lagged Analysis: experience period is claims incurred & paid within first 8 months in 2003, projection year is 2004
- 50/50 random split into Calibration & Validation

Demographic Characteristics

Demographic Category	Within Population	
	Study	Reference
Child, 0-6	5%	10%
Child, 7-22	21%	25%
Female, 23-34	4%	10%
Female, 35-49	12%	15%
Female, 50-64	24%	9%
Male, 23-34	3%	10%
Male, 35-49	10%	14%
Male, 50-64	20%	8%

- Comparison to a reference U.S. commercially insured population
- Reference used: Milliman Health Cost Guidelines, 2006 edition

Demographic Characteristics: Regional Concentration

**Table III.2 - Geographic
Characteristics**

Region	Members
Northeast	43,330
North Central	392,743
South	128,436
West	52,301
Unknown	873
Total	617,683

Calibration

- Calibration without prior cost

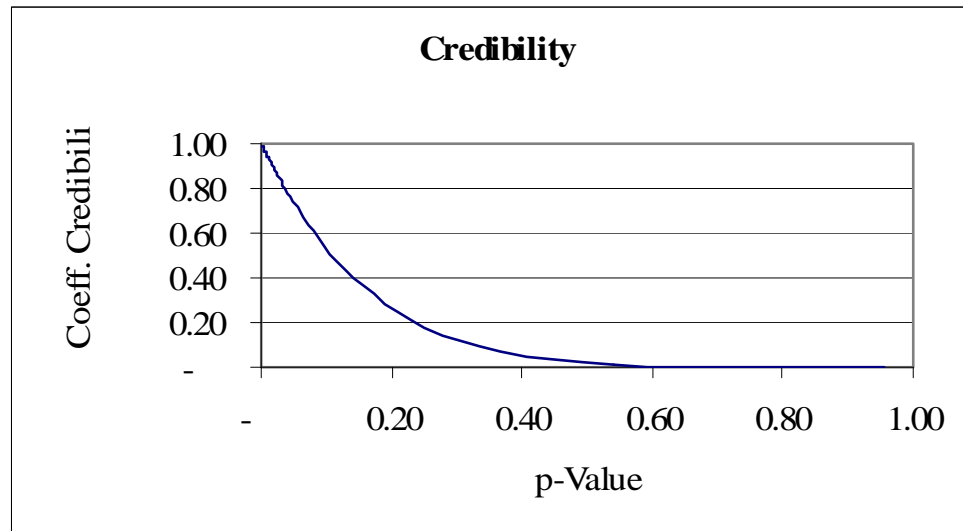
$$Y_{\text{Actual}} - Y_{\text{Prediction}} = \sum_{i=1}^A \alpha_i \times \text{Age Bin}_i + \sum_{i=1}^B \beta_i \times \text{Condition Bin}_i$$

- Calibration with prior cost

$$Y_{\text{Actual}} - Y_{\text{Prediction}} = \sum_{i=1}^A \alpha_i \times \text{Age Bin}_i + \sum_{i=1}^B \beta_i \times \text{Condition Bin}_i \\ + \gamma \times \text{Prior Cost}$$

Calibration

- Consideration of credibility (p-value criterion)



- Regressing on difference vs. cost
- Reasonability of adjustments (overall negative scores set to 0)

Results: Impact of Calibration

**Table IV.5 – R-squared and MAPE Prospective Nonlagged Offered versus Recalibrated
(Without Prior Cost, 250K Truncation)**

Risk Adjuster Tool	Inputs	R-Squared			MAPE %		
		Offered	Re-calibrated	Change	Offered	Re-calibrated	Change
ACG	Diag	19.2%	19.6%	0.4%	89.9%	88.8%	-1.1%
CDPS	Diag	14.9%	17.7%	2.8%	95.3%	91.9%	-3.4%
DxCG DCG	Diag	20.6%	21.3%	0.7%	87.5%	87.0%	-0.5%
DxCG RxGroups	Rx	20.4%	20.5%	0.1%	85.3%	85.3%	0.0%
Ingenix PRG	Rx	20.5%	21.2%	0.7%	85.8%	85.6%	-0.2%
MedicaidRx	Rx	15.8%	17.7%	1.9%	89.6%	88.4%	-1.2%
Impact Pro	Med+Rx+Use	24.4%	25.6%	1.2%	81.8%	81.6%	-0.2%
Ingenix ERG	Med+Rx	19.7%	20.0%	0.3%	86.4%	86.1%	-0.3%

We were unable to recalibrate Clinical Risk Groups*, ACG - w/ Prior Cost**, DxCG UW Model**, and MEDai**.

Censoring

- Censoring limits impact of outliers
- Censoring simulates policy limits and other risk-sharing arrangements (re-insurance)
- Error measures were adjusted to reflect the censoring, for eg. :

$$\text{R - Square} = 1 - \frac{\sum (\text{Actual}_C - \text{Predicted}_C)^2}{\sum (\text{Actual}_C - \text{Average of Actual}_C)^2},$$

Results: Impact of Censoring

Risk Adjuster Tool	Inputs	Optimized Models (Include Prior Costs)		
		250K	None	Change
ACG -w/o Prior Cost	Diag	23.0%	20.2%	2.8%
CDPS	Diag	24.6%	21.2%	3.4%
Clinical Risk Groups	Diag	20.5%	18.4%	2.1%
DxCG DCG	Diag	26.5%	22.9%	3.6%
DxCG RxGroups	Rx	27.1%	23.4%	3.7%
Ingenix PRG	Rx	27.4%	23.7%	3.7%
MedicaidRx	Rx	26.3%	22.7%	3.6%
Impact Pro	Med+Rx+Use	27.2%	24.0%	3.2%
Ingenix ERG	Med+Rx	26.5%	22.8%	3.7%
ACG - w/ Prior Cost	Diag+\$Rx	25.4%	22.1%	3.3%
DxCG UW Model	Diag+\$Total	29.1%	25.2%	3.9%
MEDai*	All	35.7%	32.1%	3.6%

With / Without Prior Cost

- Prior cost can add significant predictive value
- Using prior cost as covariate is appropriate for underwriting, not so for profiling, payment adjusting or other applications that may incent utilization
- Used aggregate prior cost: CY 2003 for non-lag, and Jan-Aug '03 for lagged analysis

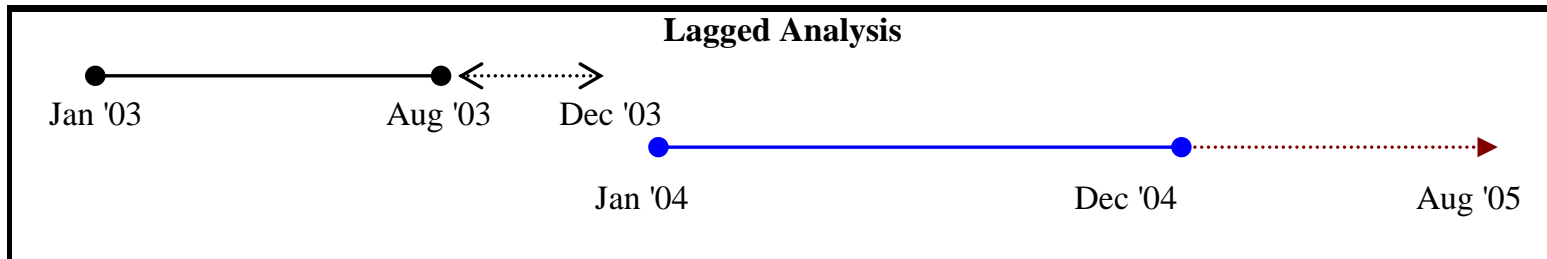
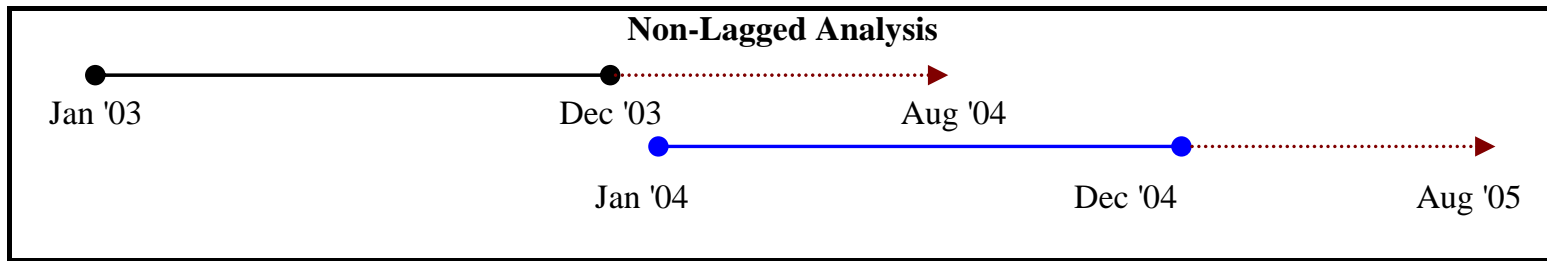
Results: With / Without Prior Costs

**Table IV.9 – R-squared and MAPE Prospective Recalibrated Nonlagged
(Without Prior Cost vs. With Prior Cost) 250K truncation**

Risk Adjuster Tool	Inputs	R-Squared			MAPE %		
		w/out Prior	with Prior	Change	w/out Prior	with Prior	Change
ACG	Diag	19.6%	23.0%	3.4%	88.8%	86.2%	-2.6%
CDPS	Diag	17.7%	24.6%	6.9%	91.9%	85.6%	-6.3%
DxCG DCG	Diag	21.3%	26.5%	5.2%	87.0%	82.5%	-4.5%
DxCG RxGroups	Rx	20.5%	27.1%	6.6%	85.3%	80.7%	-4.6%
Ingenix PRG	Rx	21.2%	27.4%	6.2%	85.6%	80.9%	-4.7%
MedicaidRx	Rx	17.7%	26.3%	8.6%	88.4%	81.9%	-6.5%
Impact Pro	Med+Rx+Use	25.6%	27.2%	1.6%	81.6%	80.6%	-1.0%
Ingenix ERG	Med+Rx	20.0%	26.5%	6.5%	86.1%	81.2%	-4.9%

Lag / Non-Lagged

- Lagged analysis follows a first renewal scenario



Results: Impact of Lag

Table IV.6 – R-squared and MAPE Offered Prospective Lagged versus Nonlagged
(Without Prior Cost) 250K Truncation

Risk Adjuster Tool	Inputs	R-Squared			MAPE %		
		Lagged	Nonlagged	Change	Lagged	Nonlagged	Change
ACG	Diag	14.5%	19.2%	4.7%	93.7%	89.9%	-3.8%
CDPS	Diag	11.9%	14.9%	3.0%	98.8%	95.3%	-3.5%
Clinical Risk Groups	Diag	14.1%	17.5%	3.4%	93.9%	90.9%	-3.0%
DxCG DCG	Diag	15.1%	20.6%	5.5%	91.6%	87.5%	-4.1%
DxCG RxGroups	Rx	18.0%	20.4%	2.4%	87.4%	85.3%	-2.1%
Ingenix PRG	Rx	18.0%	20.5%	2.5%	87.8%	85.8%	-2.0%
MedicaidRx	Rx	13.6%	15.8%	2.2%	91.7%	89.6%	-2.1%
Impact Pro	Med+Rx+Use	21.4%	24.4%	3.0%	85.5%	81.8%	-3.7%
Ingenix ERG	Med+Rx	16.9%	19.7%	2.8%	88.7%	86.4%	-2.3%

Prospective / Concurrent

- **Prospective Analysis:** using experience from CY 2003 to predict costs in CY 2004 (lag/no-lag)
- **Concurrent Analysis:** using experience from CY 2003 to predict costs in CY 2003

- **Prospective Application:** How would relative differences in morbidity look in the future?
- More uncertainty, less variation in incurred cost explained by condition indicators
- **Concurrent Application:** How did relative differences in morbidity look in the past?
- Less uncertainty, more variation in incurred cost explained by condition indicators

Results: Prospective vs. Concurrent

Table IV.8 – R-squared Offered Nonlagged (Without Prior Cost & 250K truncation) – Prospective versus

Risk Adjuster Tool	Inputs	R-Squared			MAPE %		
		Prospective	Concurrent	Change	Prospective	Concurrent	Change
ACG	Diag	19.2%	29.7%	10.5%	89.9%	75.0%	-14.9%
CDPS	Diag	14.9%	32.9%	18.0%	95.3%	80.6%	-14.7%
Clinical Risk Groups	Diag	17.5%	43.3%	25.8%	90.9%	70.5%	-20.4%
DxCG DCG	Diag	20.6%	51.8%	31.2%	87.5%	65.0%	-22.5%
MedicaidRx	Rx	15.8%	28.1%	12.3%	89.6%	79.1%	-10.5%
Ingenix ERG	Med+Rx	19.7%	42.4%	22.7%	86.4%	67.7%	-18.7%
Average		17.9%	38.0%	20.1%	89.9%	73.0%	-17.0%

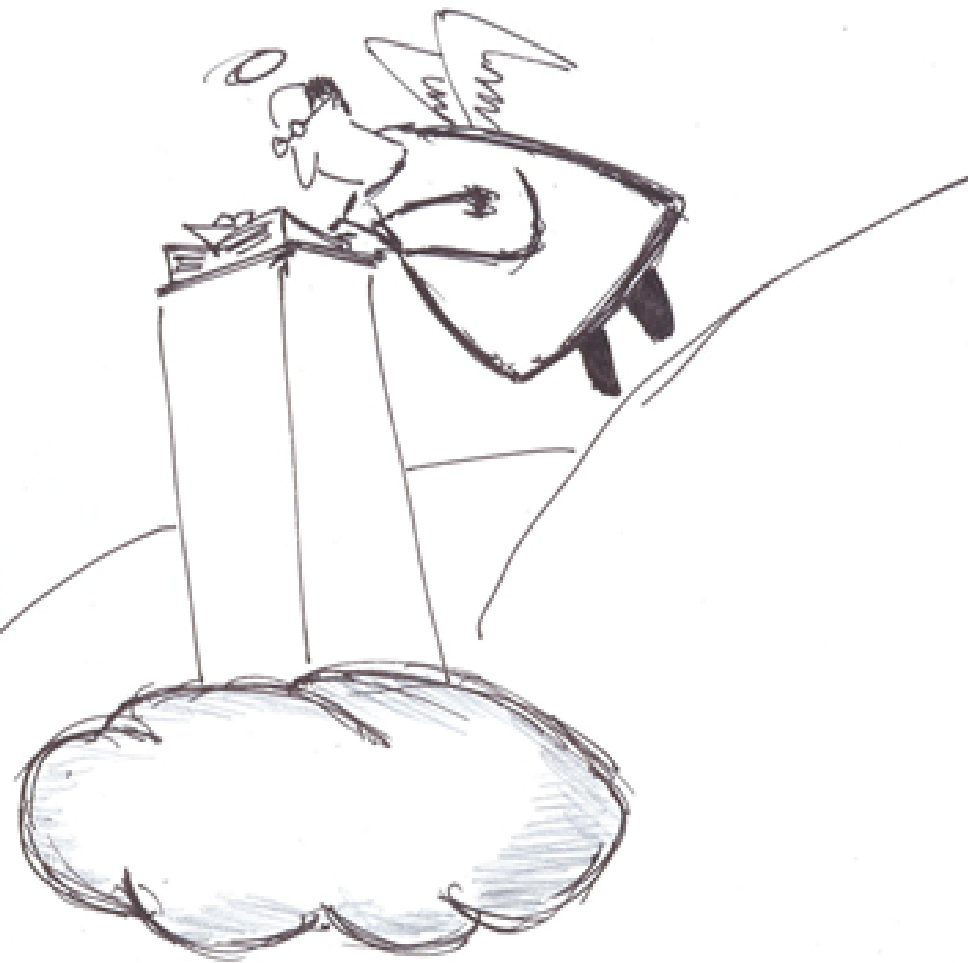
Relative Strength of Effects

	Average R-Squared Improvement
Calibration vs. Offered	0.9%
No-Lag vs. Lagged	3.3%
No Censoring vs. 250k Censor	3.4%
With Prior Costs vs. Without	5.3%
Concurrent vs. Prospective	20.1%
2002 vs 2006 Study	3.2%

Factors Affecting Predictive Performance

- Population specificity & applicability
- Turnover
- Lag Issues
- Data Issues
- Time to execute risk adjustment and frequency of updating models

But...it wasn't my time to go!



All models are wrong...
but some are useful!

Steve
04/16/04

Notes From Published Literature: A brief history of measuring performance

Commercial – Medicare – Medicaid - Veterans



Questions?

Joint Colloquium of the IACA, PBSS and IAAHS Sections
Westin Copley Place Hotel, Boston, U.S.A. – 4-7 May 2008

Further Advancements – Lifestyle Based Analytics

- Incorporating lifestyle based information in the risk adjuster
- Lifestyle Based Analytics (LBA) uses lifestyle information to enhance risk classification system for specific, relevant conditions
- Based on information from consumer data aggregators
- Use statistics and modeling techniques to find correlations between lifestyle information and prevalence of various diseases

What is Lifestyle Based Analytics?

- Medical studies have proven link between lifestyle characteristics and habits with medical conditions
- Use of lifestyle information to enhance risk classification system for specific, relevant conditions
- Based on information from consumer data aggregators
- Use statistics and modeling techniques to find correlations between lifestyle information and prevalence of various diseases

Lifestyle Based Analytics?

- In last few years, has been a lot of hype about LBA. It is important to be realistic about its power
 - 50% of acute myocardial infarctions is predicted by apoB/apoA1 ratio
 - one can hardly expect a type of vehicle owned to have more predictive power
- LBA may be used for a range of purposes
- Believe LBA will have a strong role in risk selection
 - Believe the data will improve over time
 - Believe that other forms of risk selection will be integrated with LBA

Incorporation of LBA in underwriting and pricing

- Using an automated process, name and address information passed along to data aggregator to request consumer data
 - Data available for adults ~95% of the time (varies)
- Types of information for individuals, households, neighborhood:
 - Favorite interests
 - Car choice
 - Value of home
 - Presence of high speed internet
 - ~2,200 others!

Incorporation of LBA in underwriting and pricing

- Consumer data combined with age, gender, family size
- Internal custom model receives data and calculates lifestyle adjustment factor
- Lifestyle adjustment factor centered at 1.000
- Lifestyle adjustment factor to be applied to existing base rates
- Factors typically range ~0.8 to 1.15 at group level
 - Subject to regulatory restrictions
- Within 1-2 days, rates may be delivered to prospect

Performance of LBA Predictions

- From recent client implementation:

Group Average Actual Claims Compared to Age/Gender Adjusted Average Claims	
Top 1/3 of LBA Score Members	115%
Middle 1/3 of LBA Score Members	100%
Bottom 1/3 of LBA Score Members	84%

Performance of LBA Predictions Cont'd

- **Back Pain**
 - **Occupation:** Individuals employed in mining, construction, material moving, production, installation and maintenance industries have 10% higher rate than expected based on age/sex alone
 - Individuals working in finance, management, business, etc. have an actual prevalence rate 27% lower than expected based on age/sex alone
- **Pregnancy**
 - **Family Composition:** Households with no children have 50% lower pregnancy prevalence than expected based on age/sex alone
 - **Income level:** Neighborhoods with lowest average house value has 27% higher pregnancy prevalence than expected after ruling out the effect of family composition

Performance of LBA Predictions Cont'd

- **Depression:**

Variable	Impact on Prevalence															
<ul style="list-style-type: none"> ▪ Education <ul style="list-style-type: none"> – Secondary – Post Secondary 	<p style="text-align: center;">-31%</p> <p style="text-align: center;">+25%</p>															
<ul style="list-style-type: none"> ▪ Bible / Devotional Reading <table border="0" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 33%; text-align: center;">Individual Response</td> <td style="width: 33%; text-align: center;">Neighborhood Level</td> <td style="width: 34%;"></td> </tr> <tr> <td style="text-align: center;">No</td> <td style="text-align: center;">High</td> <td style="text-align: center;">-27%</td> </tr> <tr> <td style="text-align: center;">No</td> <td style="text-align: center;">Low</td> <td style="text-align: center;">+50%</td> </tr> <tr> <td style="text-align: center;">Yes</td> <td style="text-align: center;">High</td> <td style="text-align: center;">0%</td> </tr> <tr> <td style="text-align: center;">Yes</td> <td style="text-align: center;">Low</td> <td style="text-align: center;">+51%</td> </tr> </table> 	Individual Response	Neighborhood Level		No	High	-27%	No	Low	+50%	Yes	High	0%	Yes	Low	+51%	
Individual Response	Neighborhood Level															
No	High	-27%														
No	Low	+50%														
Yes	High	0%														
Yes	Low	+51%														
<ul style="list-style-type: none"> ▪ Number of adults in household <ul style="list-style-type: none"> – 1 adult or N/A – More than 1 adult 	<p style="text-align: center;">+46%</p> <p style="text-align: center;">-2%</p>															

Development of LBA Models

- Collect data from carriers' medical claims database for one - two years
- Set aside a random portion of data to validate model after fitting
- Claims analyzed to classify people as diseased or not
- Sample diseases:

Diabetes	COPD	Cardiovascular
Hypertension	Back Pain	Depression
Arthritis	Asthma	Anxiety
Several cancers	Pregnancy	High Cholesterol
Osteoporosis		

Uses of LBA by Market Size

- **Small group market (2 to 50 employees)**
 - Some were not medically underwriting at all
 - Some carriers plan to use in conjunction with medical underwriting
 - Some carriers plan to use LBA only, without medical underwriting
- **Medium size market (51 to 250/500 employees)**
 - Medical underwriting not practical in this market
 - Experience usually not available/credible
 - Carriers often use health questionnaires (limited information)
 - LBA can be risk selection advantage against competitors

Development of LBA Models

- Use many actuarial and statistical techniques and judgment
- Combine variable values using CHAID (Chi-Squared Automatic Interaction Detection) technique to find most powerful combinations
 - Perform similar work for severity distributions, but to less effect
- Overall cost prediction divided by pure age/gender prediction
- Result: lifestyle adjustment factor for individual
 - When calculating group lifestyle adjustment factor, costs summed for all members, then divided by aggregate age/gender factor

Why the Process Works

- From pure statistical viewpoint, it doesn't matter why the correlations between lifestyle elements and disease prevalence occur – they do! (this is not a good public relations viewpoint)
- For some correlations, there is a clear cause and effect
 - More exercise -> fewer cardiovascular problems
 - Living alone -> higher risk of depression
- Sometimes, lifestyle data may reflect the condition
 - Obese people are more apt to indicate they are “walking for health”
 - Higher income members have higher rates of skin cancer
 - Driving certain cars may have greater prevalence of heart disease

Why the Process Works

- Much of the information is about neighborhood (or carrier routes)
 - Correlations regarding occupation, interests, income, wealth, cars
- Neighborhood variables provide some of our best correlations:
 - Not everyone self-reports key information about themselves
 - Knowing your neighbors have higher propensity for fitness can be a useful predictor about you
 - It is hypothesized that people in the same social network have similar health habits. Recent study: if your friends are obese, you are more likely to be as well
 - Items that seem unrelated to health show up as indicators of particular conditions
 - Rate of cultural activities correlated with COPD

Risk of Cardiovascular Disease by Favorite Interest

Reported Favorite Interest	Prevalence
Running/Jogging	0.60%
Bicycling	1.33%
Golf	2.36%
Fishing	3.02%
Camping/Hiking	3.39%
Home Workshop/Do-It-Yourself	3.42%
Avid Book Reading	3.91%
Walking for Health	4.03%
Flower Gardening	4.55%

Summary of LBA

- LBA is an additive tool to enhance risk selection
- Does not pick out specific members in group as definitely having particular condition
- Does identify meaningful differences from one person to another, and from one group to another
- Will give advantages to early adopters

Questions?

Thank you!