SCIOinspire
FRESH thinking REAL results

Innovating Healthcare Business Process Service Delivery

Predictive Modeling Basics and Beyond

September, 2008

Agenda

1. Background and Issues.
2. Model Objectives.
3. Population/Data.
4. Sample model.
5. Program Planning.
6. Program Evaluation.
7. Risk Transition.
8. General discussion.

Introductions

Ian Duncan, FSA, FIA, FCIA, MAAA.

President, Solucia Consulting, a SCIOinspire Company.

Kate Hall, ASA MAAA

Vice President, Solucia Consulting, a SCIOinspire Company.

Introduction / Objective

1. What is Predictive Modeling?
2. Types of predictive models.
3. Data and Data Preparation.
4. Applications – case studies.

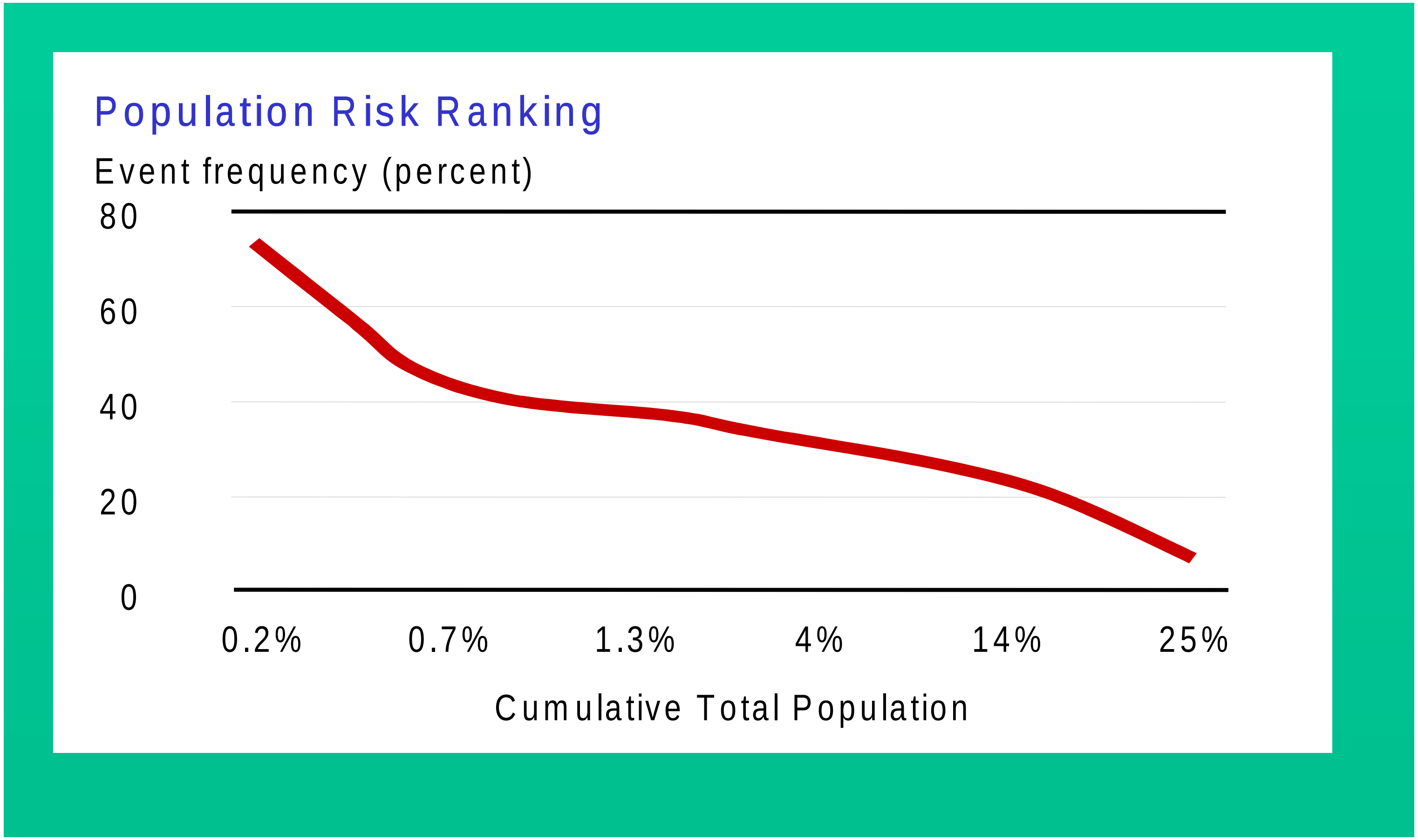
Predictive Modeling:
A Review of the Basics

Definition of Predictive Modeling

"Predictive modeling is a set of tools used to stratify a population according to its risk of nearly any outcome…ideally, patients are risk-stratified to identify opportunities for intervention <u>before</u> the occurrence of adverse outcomes that result in increased medical costs."

Cousins MS, Shickle LM, Bander JA. An introduction to predictive modeling for disease management risk stratification. Disease Management 2002;5:157-167.

“Stratified according to risk of event”



PM – more often wrong than right...

"The year 1930, as a whole, should prove at least a fairly good year."

-- *Harvard Economic Service, December 1929*

Why do it? Potential Use of Models

Program Management Perspective

- Identifying individuals at very high risk of an event (death, LTC, disability, annuity surrender, etc.).
- Identify management opportunities and determine resource allocation/ prioritization.

SCIOinspire
FRESH thinking REAL results

Identification – how?

- At the heart of predictive modeling!
  - Who?
  - What common characteristics?
  - What are the implications of those characteristics?
- There are many different algorithms for identifying member conditions. THERE IS NO SINGLE AGREED FORMULA.
- Condition identification often requires careful balancing of sensitivity and specificity.

A word about codes and groupers

Codes are the “raw material” of predictive modeling.

Codes are required for payment, so they tend to be reasonably accurate - providers have a vested interest in their accuracy.

Codes define important variables like Diagnosis (ICD-9 or 10); Procedure (CPT); Diagnosis Group (DRG – Hospital); Drug type/dose/manufacturer (NDC); lab test (LOINC); Place of service, type of provider, etc. etc.

“Grouper” models sort-through the raw material and consolidate it into manageable like categories.

Identification – example (Diabetes)

Diabetics can be identified in different ways:

Diagnosis type	Reliability	Practicality
Physician Referral	High	Low
Lab tests	High	Low
Claims	Medium	High
Prescription Drugs	Medium	High
Self-reported	Low/medium	Low

Medical and Drug Claims are often the most practical method of identifying candidates for predictive modeling.

Identification – example (Diabetes)

Inpatient Hospital Claims – ICD-9 Claims Codes

ICD-9-CM CODE DIABETES	DESCRIPTION
250.xx	Diabetes mellitus
357.2	Polyneuropathy in diabetes
362.0, 362.0x	Diabetic retinopathy
366.41	Diabetic cataract
648.00-648.04	Diabetes mellitus (as other current condiition in mother classifiable elsewhere, but complicating pregnancy, childbirth or the puerperioum.

Less severe

More Severe

Diabetes – additional codes

CODES	CODE TYPE	DESCRIPTION - ADDITIONAL
DIABETES;		
G0108, G0109	HCPCS	Diabetic outpatient self-management training services, individual or group
J1815	HCPCS	Insulin injection, per 5 units
67227	CPT4	Destruction of extensive or progressive retinopathy, (e.g. diabetic retinopathy) one or more sessions, cryotherapy, diathermy
67228	CPT4	Destruction of extensiive or progressive retinopathy, one or more sessions, photocoagulation (laser or xenon arc).
996.57	ICD-9-CM	Mechanical complications, due to insulin pump
V45.85	ICD-9-CM	Insulin pump status
V53.91	ICD-9-CM	Fitting/adjustment of insulin pump, insulin pump titration
V65.46	ICD-9-CM	Encounter for insulin pump training

Diabetes – drug codes

Insulin or Oral Hypoglycemic Agents are often used to identify members. A simple example follows; for more detail, see the HEDIS code-set.

This approach is probably fine for Diabetes, but may not work for other conditions where off-label use is prevalent.

Insulin	
2710*	Insulin**

OralAntiDiabetics	
2720*	Sulfonylureas**
2723*	Antidiabetic - Amino Acid Derivatives**
2725*	Biguanides**
2728*	Meglitinide Analogues**
2730*	Diabetic Other**
2740*	ReductaseInhibitors**
2750*	Alpha-Glucosidase Inhibitors**
2760*	Insulin Sensitizing Agents**
2799*	Antiadiabetic Combinations**

SCIOinspire
FRESH thinking REAL results

More about Grouper Models

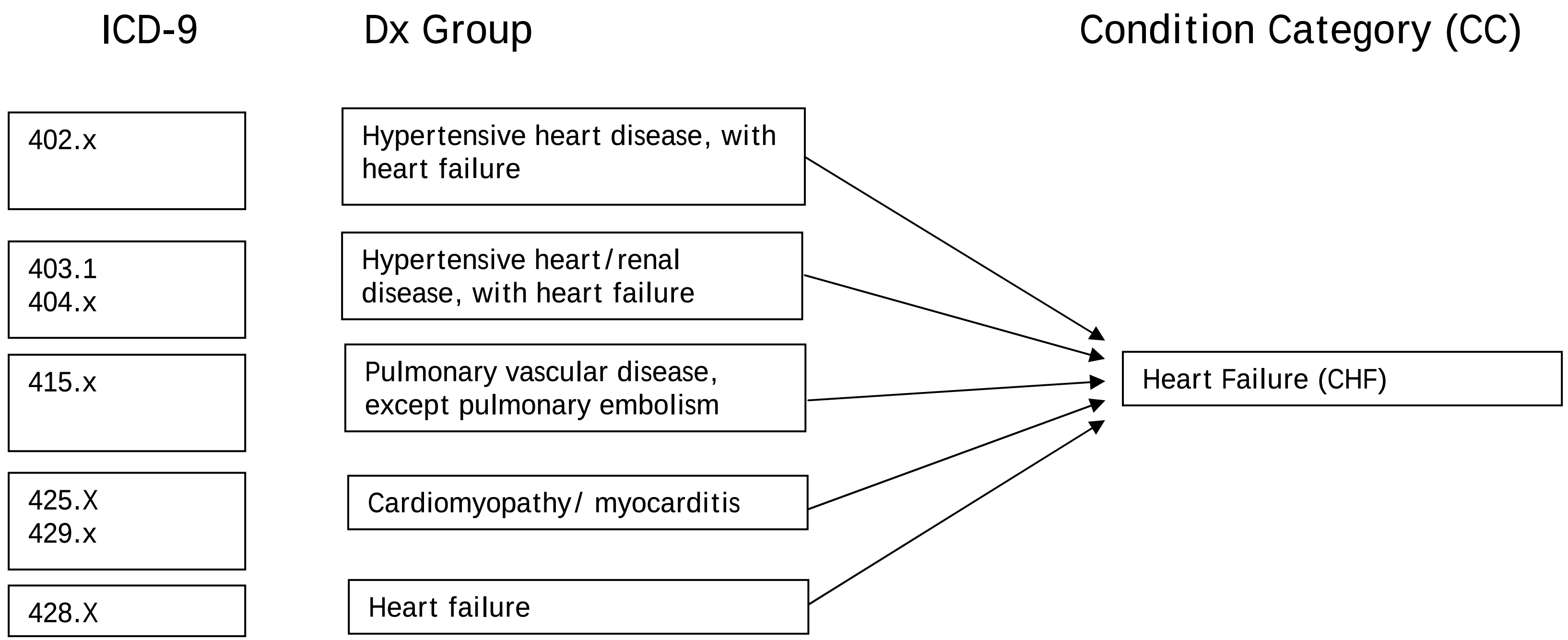
Grouper models address several problems inherent in identification from claims (medical and/or drug):

- What "recipe" or algorithm to apply?
- How to keep the algorithm up-to-date?
- How to achieve consistency among users (important, for example, in physician reimbursement or program assessment).

They also have draw-backs:

- Someone else's definitions;
- Lack of transparency;
- You can't control sensitivity/specificity trade-off.

Grouper Models – example



- Each Group and Condition Category becomes an independent variable in a multiple regression equation that results in a weight for that condition;
- Weights correlate with average resource utilization for that condition;
- Some are "trumped" by others (more severe);
- Scores can range from ≅ 0.0 (for young people without diagnoses) to numbers in the 40's and 50's (for multiple co-morbid patients).

Construction of a model*

* From Ian Duncan: "Managing and Evaluating Care Management Interventions" (Actex, 2008)

Condition Category	Risk Score Contribution	Notes
Diabetes with No or Unspecified Complications	0.0	Trumped by Diabetes with Renal Manifestation
Diabetes with Renal Manifestation	2.1	
Hypertension	0.0	Trumped by CHF
Congestive Heart Failure (CHF)	1.5	
Drug/Alcohol Dependence	0.6	
Age-Sex	0.4	
Total Risk Score	**4.6**	

Construction of a model

Grouper/Risk-adjustment theory is that there is a high correlation between risk scores and actual dollars (resources used).

The Society of Actuaries has published three studies that test this correlation. They are available from the SOA and are well worth reading. (See bibliography.) They explain some of the theory of risk-adjusters and their evaluation, as well as showing the correlation between $'s and Risk Scores for a number of commercial models.

Note 1: the SOA tests both *Concurrent* (retrospective) and *Prospective* models. Concurrent model correlations tend to be higher.

Note 2: there are some issues with models that you should be aware of:

- They tend to be less accurate at the "extremes" (members with high or low risk scores);
- We have observed an inverse correlation between risk-score and $'s across a wide range of members.

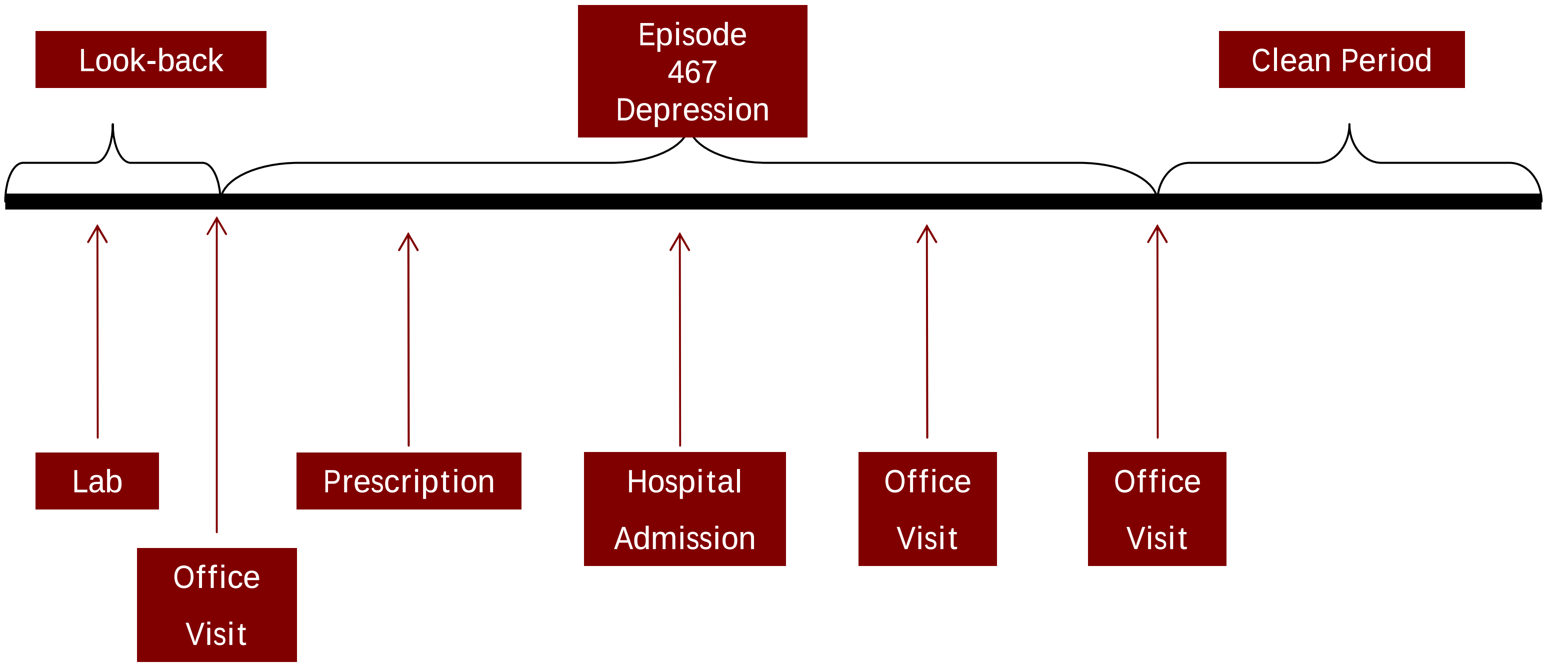
A different approach to grouping

Grouping by Episode

Services related to the underlying diagnosis are grouped

Different diagnosis related groups have different cost weights.

Complete/Incomplete groups



All people are not equally identifiable

Definition Examples:

Narrow:	Hospital Inpatient (primary Dx); Face-to-face professional (no X-Ray or Lab)
Broad:	Hospital I/P (any Dx); All professional including X-ray, lab.
Rx:	Narrow + Outpatient Prescription

Prevalence of 5 Chronic conditions

	Narrow	Broad	Rx
Medicare	24.4%	32.8%	30.8%
Commercial	4.7%	6.3%	6.6%

SCIOinspire
FRESH thinking REAL results

Identification: False Positives/ False Negatives

False Positive Identification Incidence through Claims
Medicare Advantage Population (with drug benefits)
Diabetes Example

		Year 1 Narrow	Year 1 + Broad	Year 1 + Rx	TOTAL
Year 2	Narrow	75.9%			
	+ Broad		85.5%		
	+ Rx			92.6%	
	Not Identified	24.1%	14.5%	7.4%	
	TOTAL	100.0%	100.0%	100.0%	100.0%

Why do it? Model Applications

Reimbursement

- Predicting (normalized) resource use in the population.

Example 1: Normalized resources

Remember the "Scores" we introduced a few slides back?

PROVIDER GROUP XXX

Member Group ID	Condition(s)	# members	Score	Risk Total	Expected Cost	Actual Cost
1080	CHF	2	19.9	39.8	$ 43,780	$ 50,000
532	Cancer 1	20	8.7	174.2	$ 191,620	$ 150,000
796	Cancer 2 + Chronic condition	10	16.0	159.7	$ 175,670	$ 160,000
531	Cancer 2 + No chronic condition	15	9.0	135.3	$ 148,830	$ 170,000
1221	Multiple chronic conditions	6	4.8	28.8	$ 31,680	$ 50,000
710	Acute + Chronic Conditions	10	11.1	110.9	$ 121,990	$ 125,000
882	Diabetes	7	3.7	25.7	$ 28,270	$ 28,000
967	Cardiac	4	6.1	24.5	$ 26,950	$ 30,000
881	Asthma	8	3.0	24.1	$ 26,510	$ 40,000
		82		723.0	$ 795,300	$ 803,000

SCIOinspire
FRESH thinking REAL results

Why do it? Model Applications

Program Evaluation

- Predicting resource use based on condition profile.
- Trend Adjustment.

Example 2: Program Evaluation

Typical Program Evaluation Methodology (e.g. DMAA)

Estimated Savings due to reduced PMPY =

Baseline Cost PMPY $\times$ Cost Trend		$\$6{,}000 \times 1.12 = \$6{,}720$
Minus:	Actual Cost PMPY	$\underline{\$6{,}300}$
Equals:	Reduced Cost PMPY	$420
Multiplied by:	Actual Member Years in Measurement Period	$\underline{20{,}000}$
Estimated Savings		$8,400,000

Trend can be biased by changes in population risk-profile over time; adjustment for change in average risk will correct for this.

Why do it? Potential Uses of Models

Actuarial, Underwriting

- Calculating new business and renewal premiums

Example 3: New/Renewal premium

Kate Hall will be speaking about this issue later.

Why do it? Potential Uses of Models

Provider Profiling

- Profiling of provider
- Efficiency Evaluation
- Provider & health plan contracting

Example 4: Provider profiling

Different approaches: provider panel resource prediction (example 1) OR Episode Risk projection

Efficiency Assessment

Current time period: Jan 2004 - Dec 2004

Previous time period: Jan 2003 - Dec 2003

Applied Filter Criteria:

Advanced Sort + | Filter +

Rows 1..10 of 10

Provider Group	Patient Complexity Index	Current Patients	Current Episodes	Current Actual Average Allowed Amount per Episode	Current Expected Average Allowed Amount per Episode	Current Performance Ratio Actual to Expected	Previous Patients	Previous Episodes	Previous Actual Average Allowed Amount per Episode	Previous Expected Average Allowed Amount per Episode	Previous Performance Ratio Actual to Expected	Percentage Trend Current Year to Previous Year
790518738	101.20	113	127	$1,050	$1,089	1.0	108	123	$673	$703	1.0	-1 %
294031519	98.77	50	53	$637	$1,063	0.6	50	58	$452	$881	0.5	-17%
980070973	100.10	47	64	$1,523	$1,077	1.4	40	51	$1,581	$930	1.7	17%
415326972	103.44	42	48	$1,445	$1,113	1.3	37	43	$1,060	$1,014	1.0	-24 %
724236523	107.99	42	43	$745	$1,162	0.6	35	38	$687	$876	0.8	18 %
739159523	77.66	37	49	$749	$835	0.9	39	55	$474	$678	0.7	-28 %

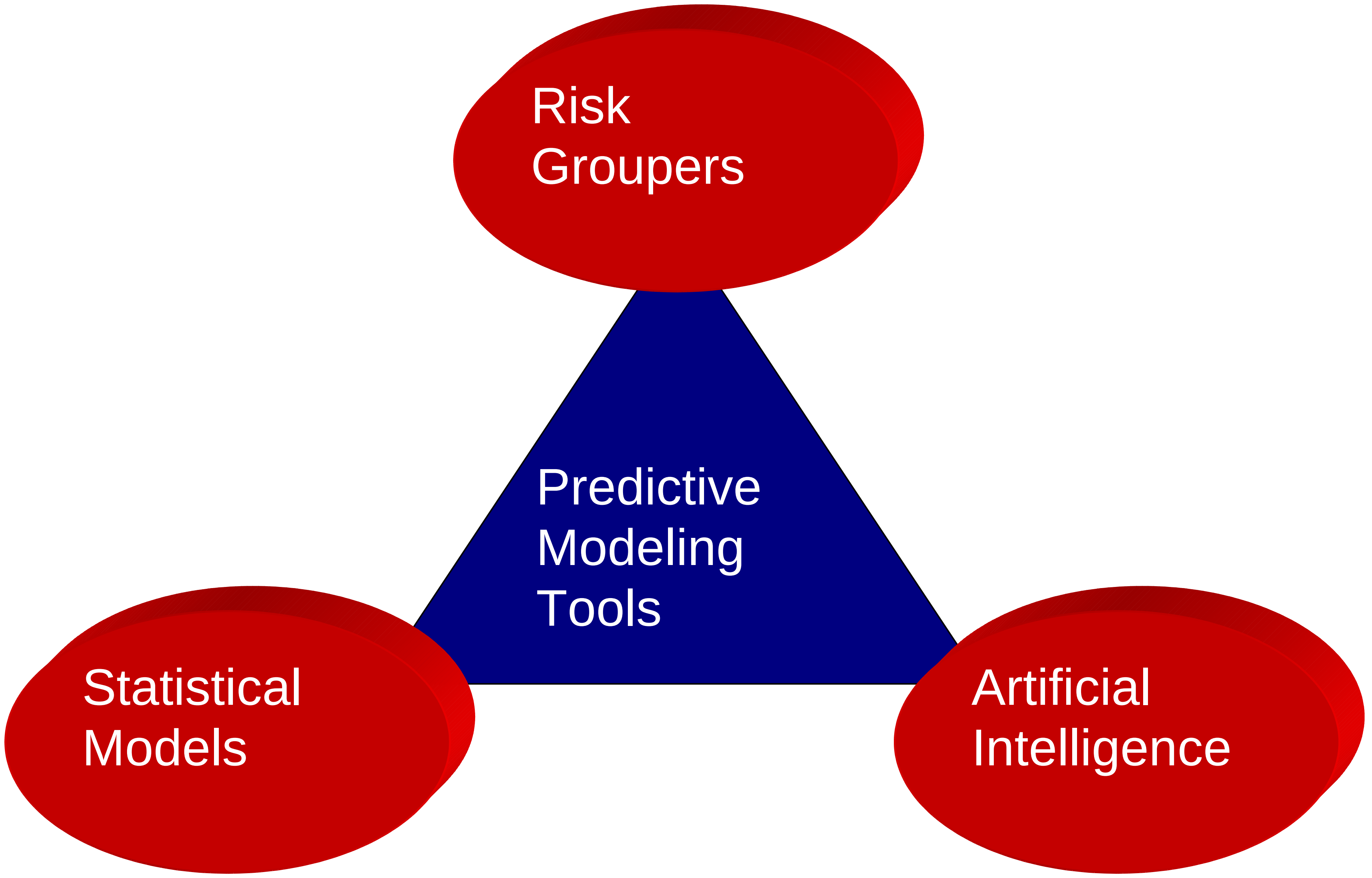
SCIOinspire FRESH thinking REAL results

Why do it? Potential Uses of Models

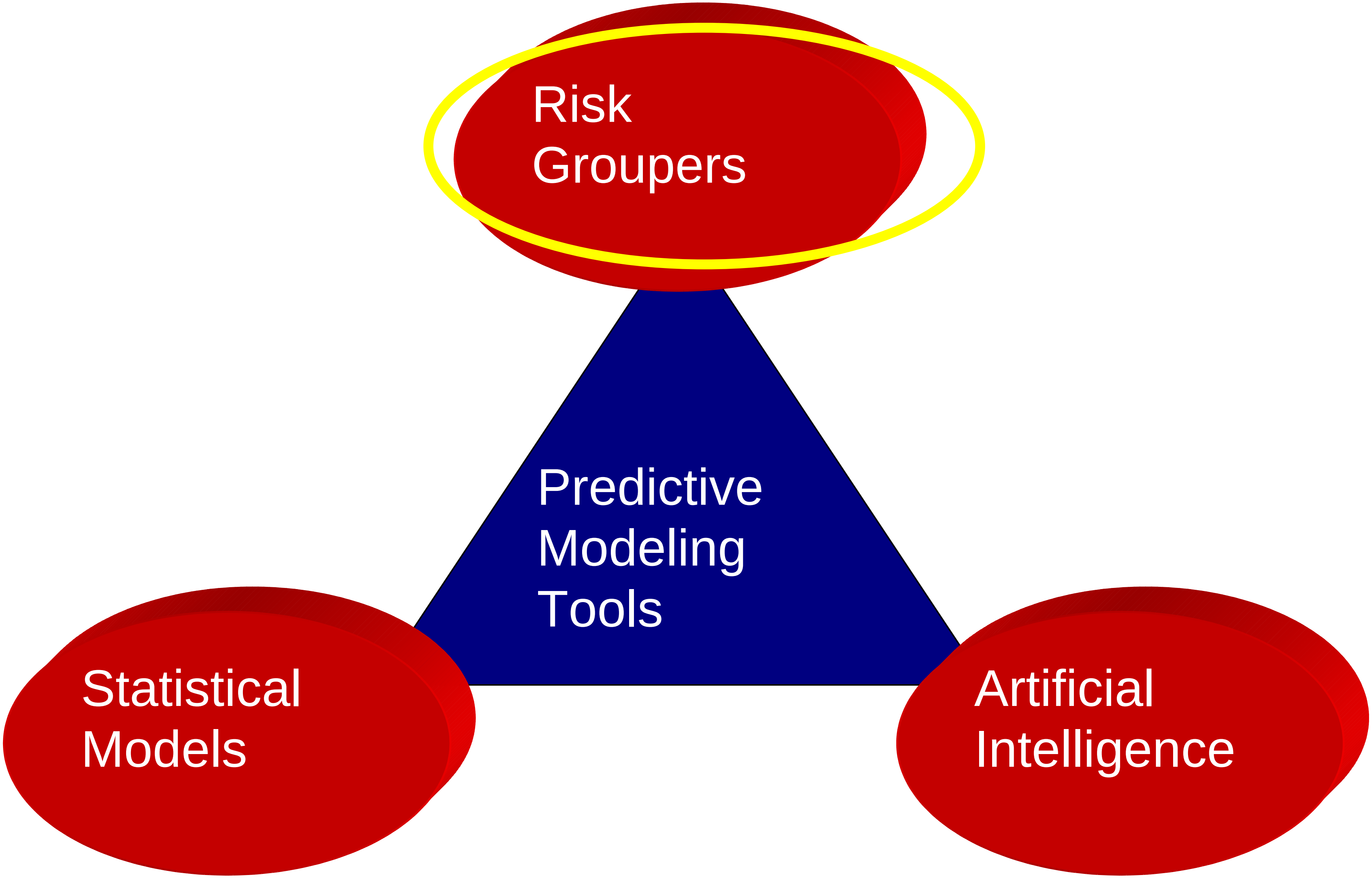
From a Medical Management Perspective

- Identifying individuals at very high risk for high utilization
- Resource allocation and program planning.

Types of Predictive Modeling Tools



Types of Predictive Modeling Tools



Uses of Risk Groupers

Actuarial, Underwriting and Profiling Perspectives

Medical Management Perspective

Program Evaluation Perspective

Risk Groupers

What are the different types of risk groupers?

Selected Risk Groupers

Company	Risk Grouper	Data Source
IHCIS/Ingenix	ERG	Age/Gender, ICD-9 NDC, Lab
UC San Diego	CDPS	Age/Gender, ICD -9 NDC
DxCG	DCG RxGroup	Age/Gender, ICD -9 Age/Gender, NDC
Symmetry/Ingenix	ERG PRG	ICD – 9, NDC NDC
Johns Hopkins	ACG	Age/Gender, ICD – 9

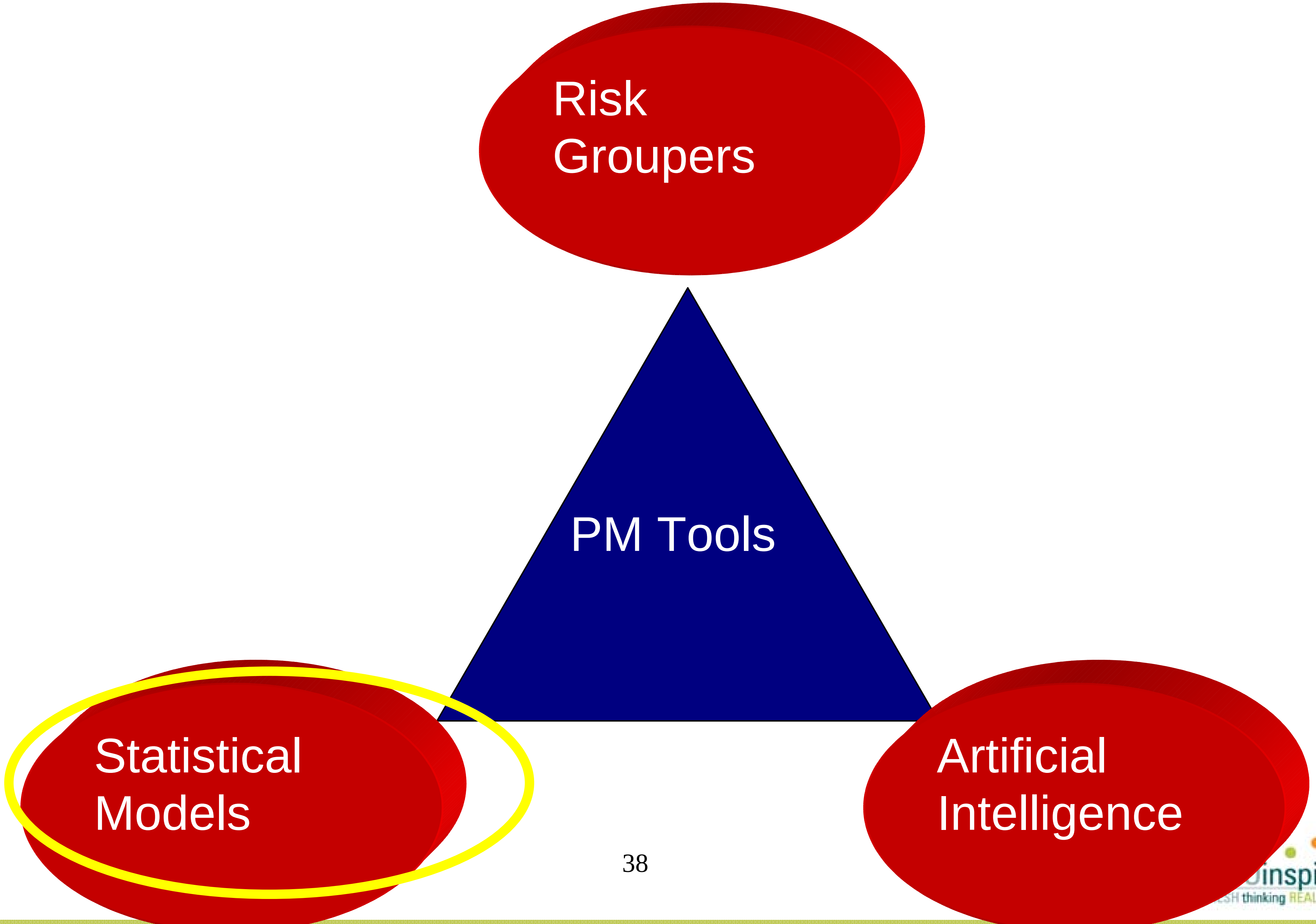
SCIOinspire
FRESH thinking REAL results

Risk Grouper Summary

1. Similar performance among all leading risk groupers*.
2. Risk grouper modeling tools use *different algorithms* to group the source data.
3. Risk groupers use *relatively limited data* sources (e.g. DCG and Rx Group use ICD-9 and NDC codes but not lab results or HRA information)
4. Most Risk Grouper based Predictive Models combine also use statistical analysis.

* See New SOA study (Winkelman et al) published 2007. Available from SOA.

Types of Predictive Modeling Tools



Uses of Statistical Models

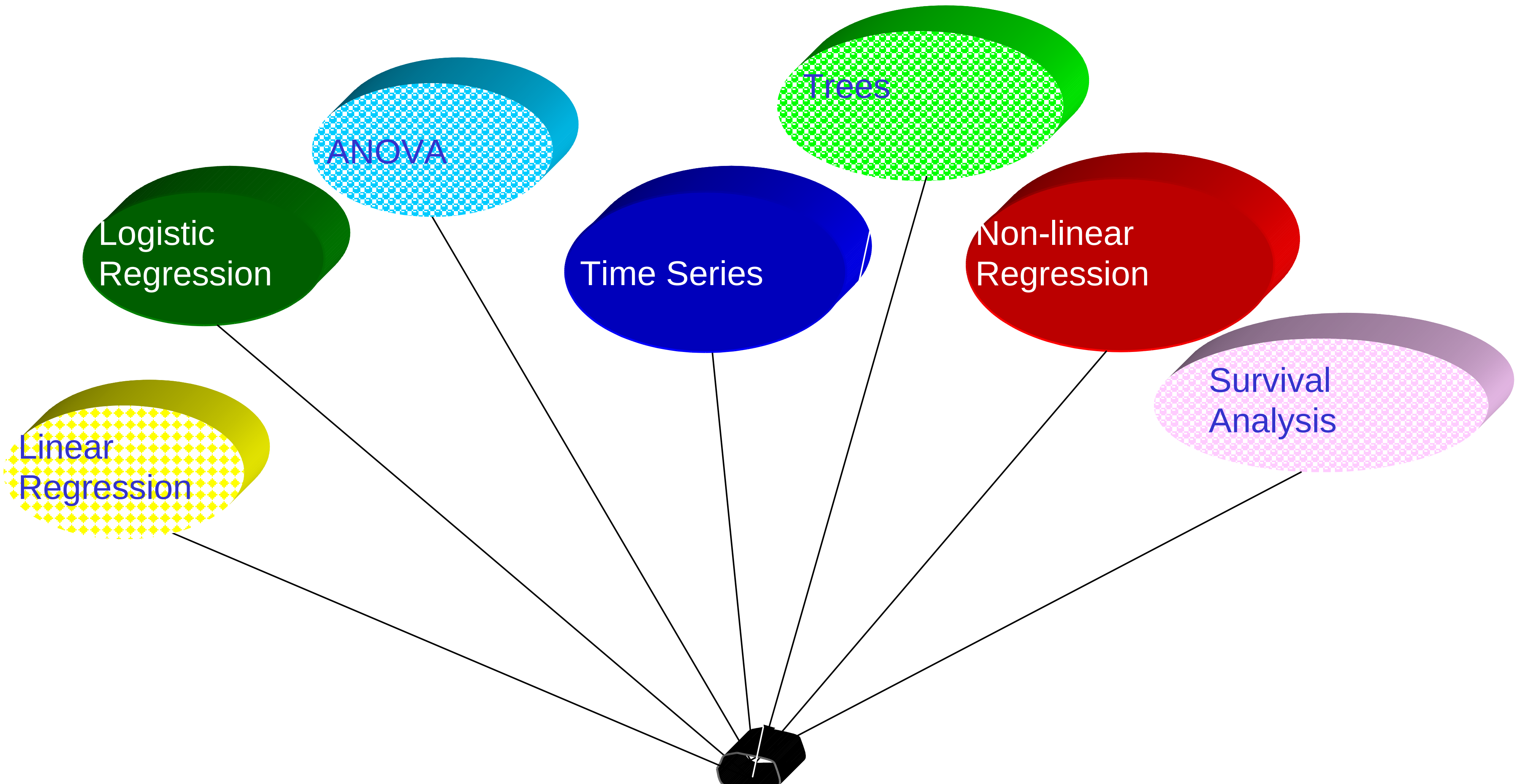
Statistical models can be used for all 3 uses

Medical Management Perspective	Actuarial, Underwriting and Profiling Perspectives	Program Evaluation Perspective

Statistical Models

What are the different types of statistical models?

Types of Statistical Models



Statistical Models

Time series modeling tools is another type of statistical modeling tool – it requires a lot of historical data.

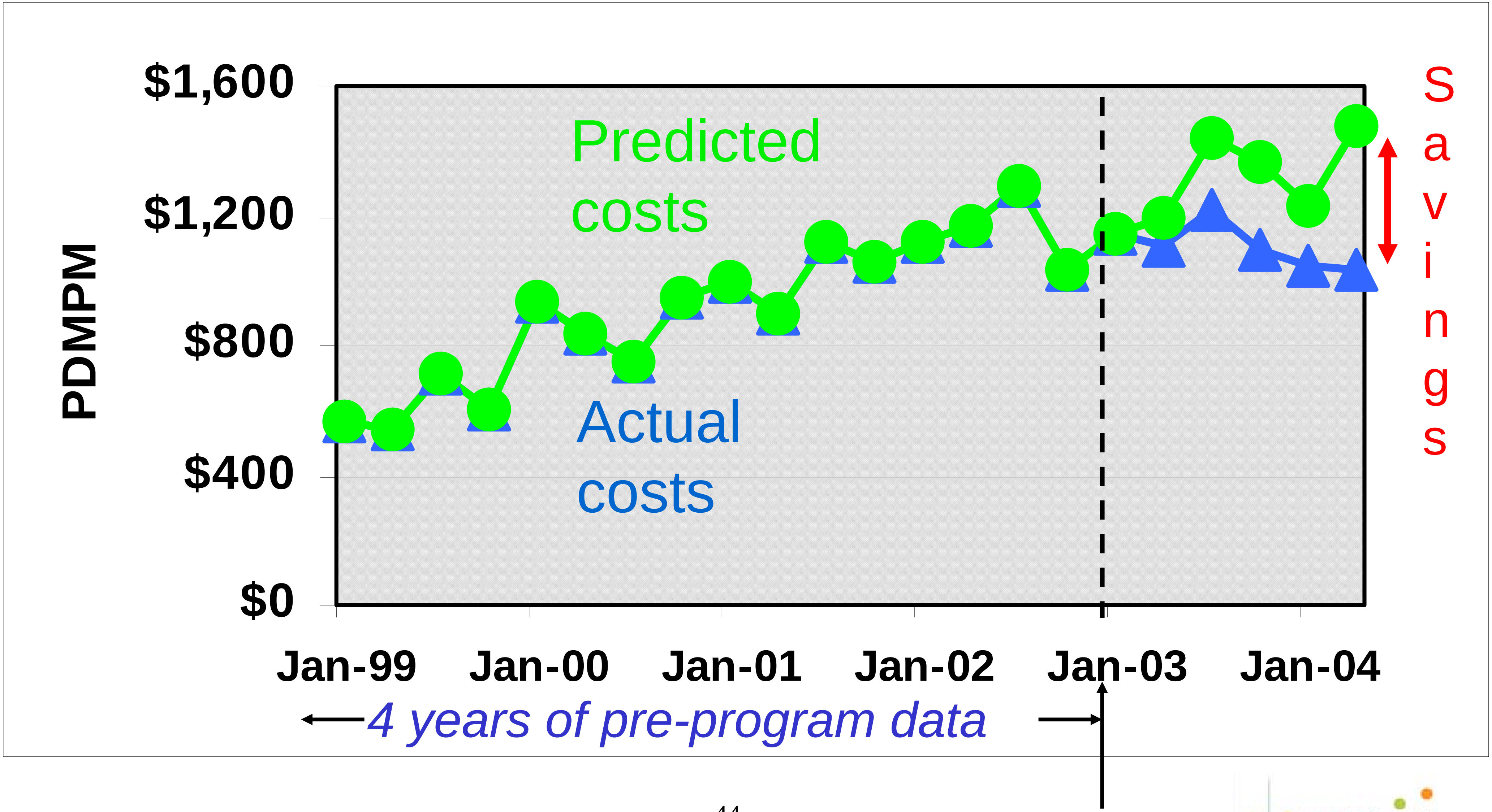
Time Series

Time series analysis is to

a) Identify the pattern of observed time series data and

b) Forecast future values by extrapolating the identified pattern.

Example: Time Series

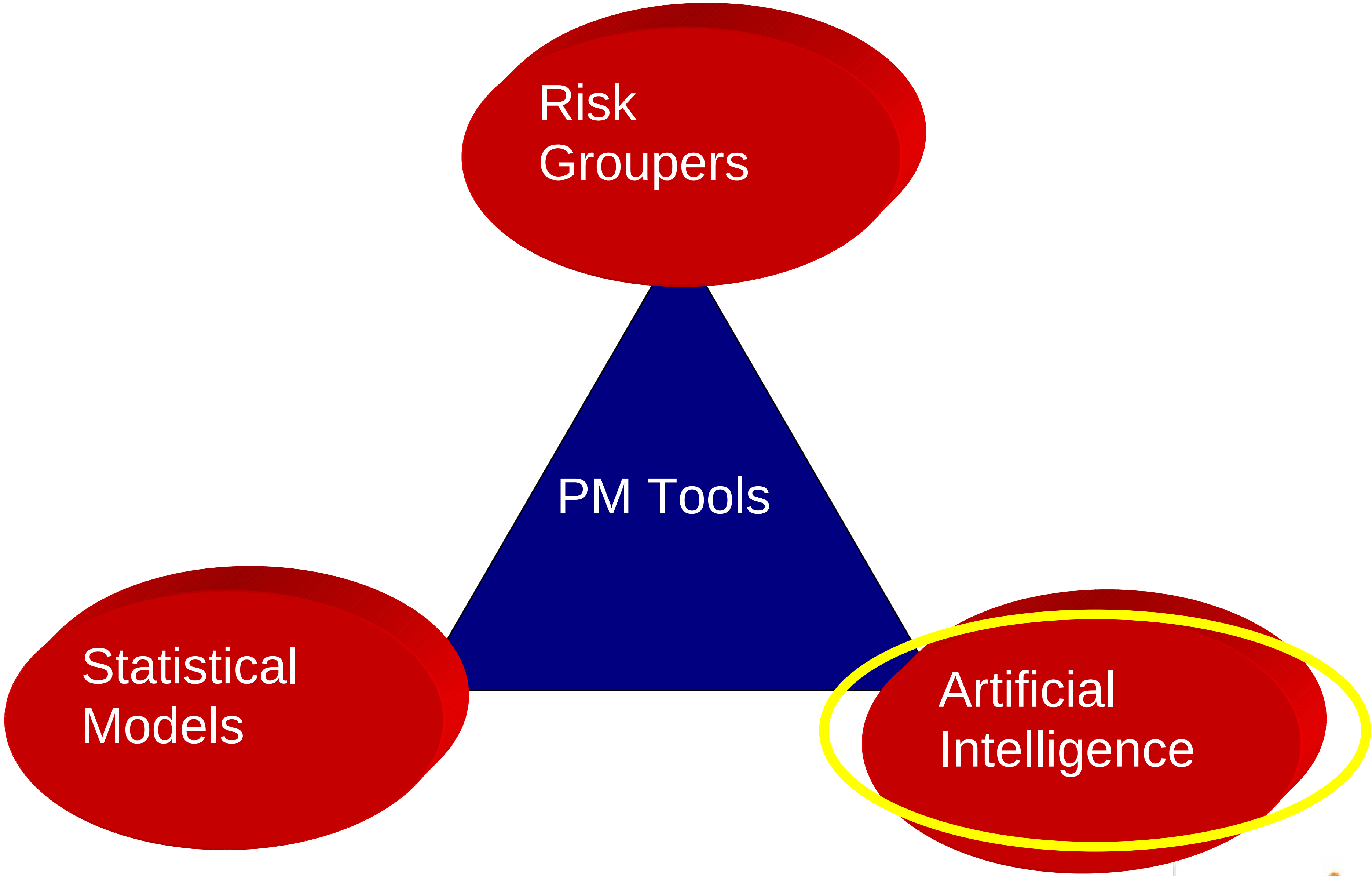


Statistical Model Summary

1. Statistical models can be used for a number of actuarial applications: evaluation, premium calculation, provider profiling and resource allocation.
2. The predictive model is a critical component of successful medical management intervention programs -“impactability is key in medical management”.
3. Statistical models can use all available detailed data (e.g. lab results or HRA).

SCIOinspire FRESH thinking REAL results

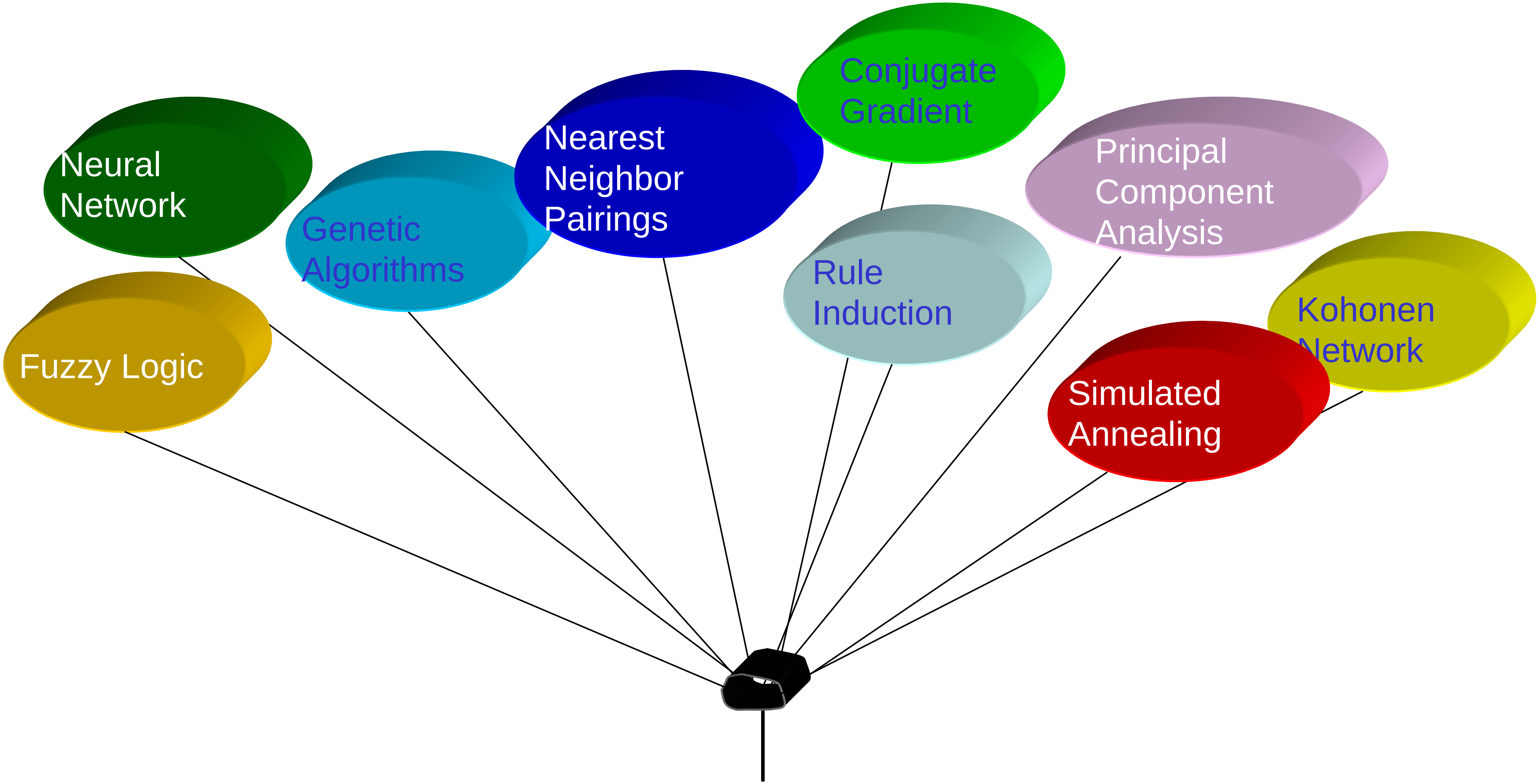
Types of Predictive Modeling Tools



Artificial Intelligence Models

What are the different types of artificial intelligence models?

Artificial Intelligence Models



Features of Neural Networks

Reality

NN tracks complex relationships by resembling the human brain

Perception

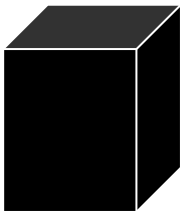
NN can accurately model complicated health care systems

Reality

- Performance equals standard statistical models
- Models overfit data

Neural Network Summary

1. Good academic approach.
2. Few data limitations.
3. Performance comparable to other approaches.
4. Can be hard to understand the output of neural networks (black box).



In Summary

1. Leading predictive modeling tools have similar performance.
2. Selecting a predictive modeling tool should be based on your specific objectives - one size doesn’t fit all.
3. A good predictive model for medical management should be linked to the intervention (e.g. impactibility).
4. “Mixed” models can increase the power of a single model.

Rules vs. Prediction

We are often asked about rules-based models.

1. First, all models ultimately have to be converted to rules in an operational setting.
2. What most people mean by "rules-based models" is actually a "Delphi*" approach. For example, application of "Gaps-in-care" or clinical rules (e.g. ActiveHealth).
3. Rules-based models have their place in Medical Management. One challenge, however, is risk-ranking identified targets, particularly when combined with statistical models.

* Meaning that experts determine the risk factors, rather than statistics.

PM is <u>NOT</u> always about *Cost Prediction…..*

…..it IS about resource allocation.

- Where/how should you allocate resources?
- Who is *intervenable* or *impactable*?
- What can you expect for outcomes?
- How can you manage the key drivers of the economic model for better outcomes?

Cost Stratification of a Large Population

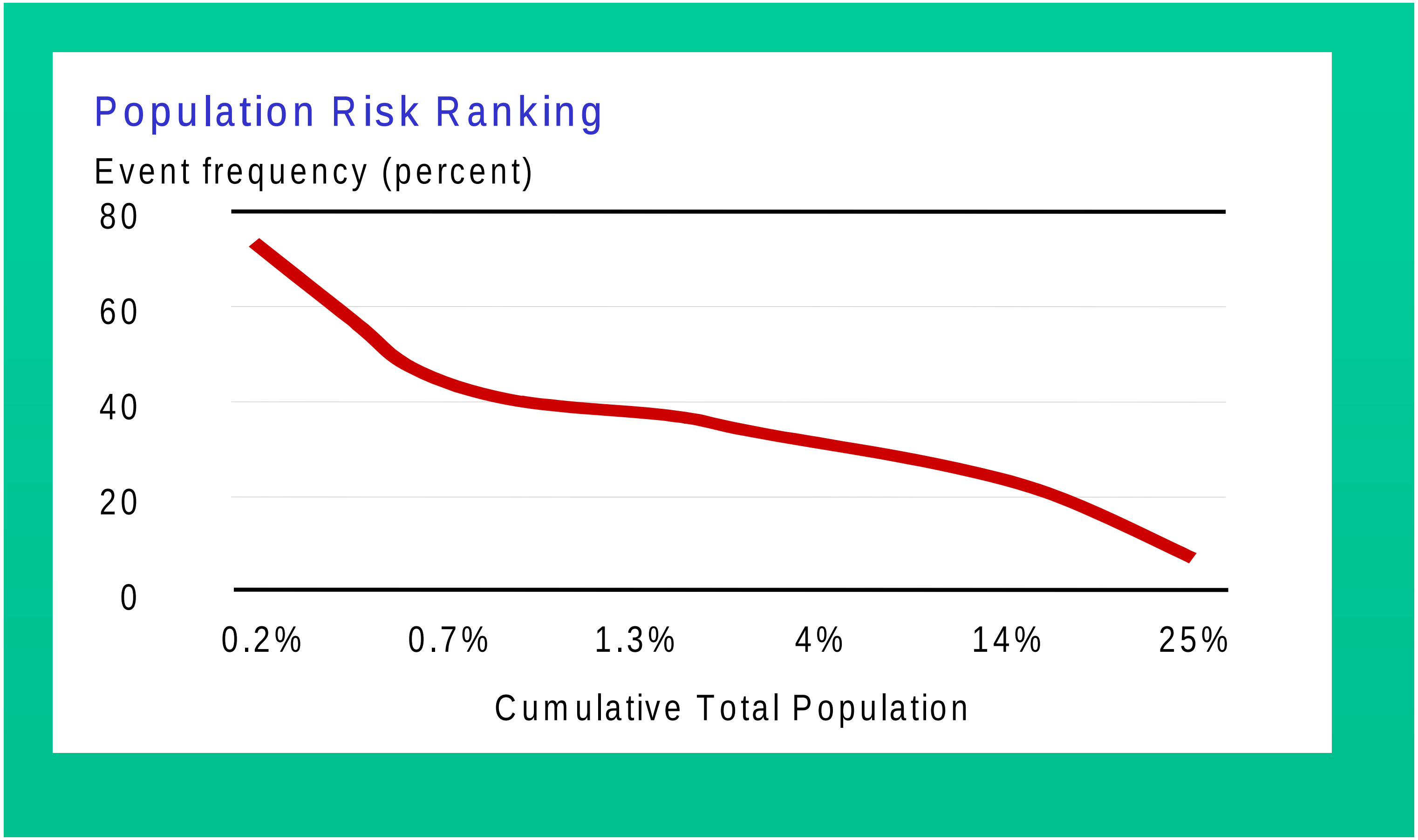
	0.0% - 0.5%	0.5% - 1.0%	Top 1%	Top 5%	Total
Population	67,665	67,665	135,330	676,842	13,537,618
Actual Cost	$3,204,433,934	$1,419,803,787	$4,624,237,721	$9,680,579,981	$21,973,586,008
PMPY Total Actual Cost	$47,357	$20,977	$34,170	$14,303	$1,623
Percentage of Total Cost	**14.6%**	**6.5%**	**21.1%**	44.1%	100%
Patients with > $50,000 in Claims					
	0.0% - 0.5%	**0.5% - 1.0%**	**Top 1%**	**Top 5%**	**Total**
Number of Patients	19,370	5,249	24,619	32,496	35,150
Percentage of Total	55.1%	14.9%	**70.0%**	92.4%	100.0%

Cost Stratification of a Large Population

	0.0% - 0.5%	0.5% - 1.0%	Top 1%	Top 5%	Total
Population	67,665	67,665	135,330	676,842	13,537,618
Actual Cost	$3,204,433,934	$1,419,803,787	$4,624,237,721	$9,680,579,981	$21,973,586,008
PMPY Total Actual Cost	$47,357	$20,977	$34,170	$14,303	$1,623
Percentage of Total Cost	**14.6%**	**6.5%**	**21.1%**	44.1%	100%
Patients with > $50,000 in Claims					
	0.0% - 0.5%	**0.5% - 1.0%**	**Top 1%**	**Top 5%**	**Total**
Number of Patients	19,370	5,249	24,619	32,496	35,150
Percentage of Total	55.1%	14.9%	**70.0%**	92.4%	100.0%

SCIOinspire
FRESH thinking REAL results

Decreasing Cost / Decreasing Opportunity



Important Concept: this chart represents *Predicted*, not *Actual* Cost.

The Economic Model and Planning a Program

- As the Population Risk Ranking slide shows, all people do not represent equal opportunity.
- The difference in opportunity means that programs need to be well planned.
- It also gives you an opportunity to test the accuracy of different models.

Economic Model: Simple example

- 30,000 eligible members (ee/dep)
- 1,500 – 2,000 with chronic conditions
- 20% “high risk” – 300 to 400
- 60% are reachable and enroll: 180 - 240
- Admissions/high-risk member/year: 0.65
- “Change behavior” of 25% of these:

  - reduced admissions: 29 to 39 annually

  - cost: $8,000/admission

- Gross Savings: $232,000 to $312,000

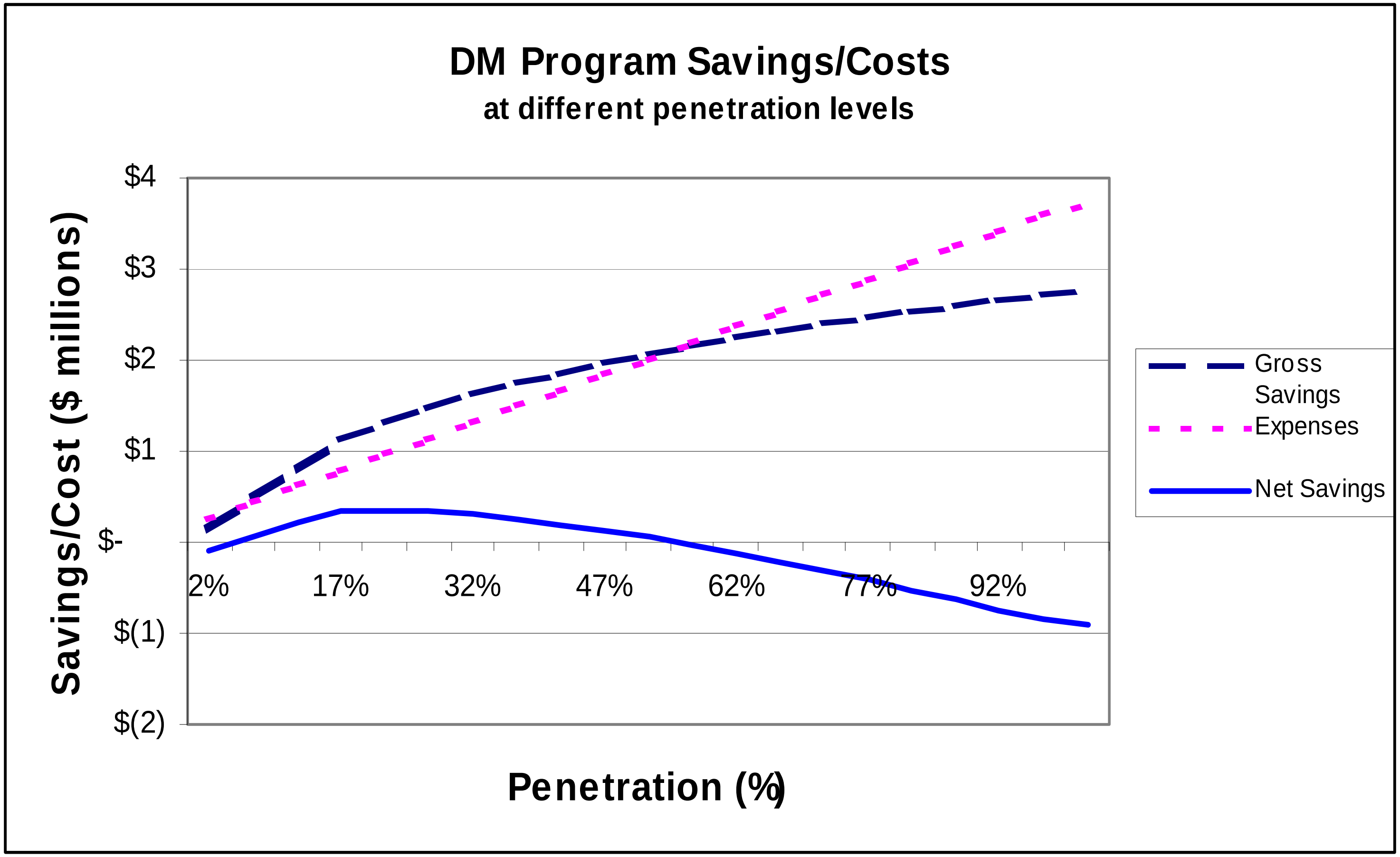
  - $0.64 to $0.87 pmpm.

SCIOinspire
FRESH thinking REAL results

Key drivers of the economic model

- Prevalence within the population (numbers)
- Ability to Risk Rank the Population
- Data quality
- Reach/engage ability
- Cost/benefit of interventions
- Timeliness
- Resource productivity
- Random variability in outcomes

Understanding the Economics



Practical Example of Model-Building

What is a model?

- A model is a set of coefficients to be applied to production data in a live environment.
- With individual data, the result is often a predicted value or "score". For example, the likelihood that an individual will purchase something, or will experience a high-risk event (surrender; claim, etc.).
- For underwriting, we can predict either cost or risk-score.

Background

Available data for creating the score included the following

- Eligibility/demographics
- Rx claims
- Medical claims

For this project, several data mining techniques were considered: neural net, CHAID decision tree, and regression. The regression was chosen for the following reasons:

With proper data selection and transformation, the regression was very effective, more so than the tree.

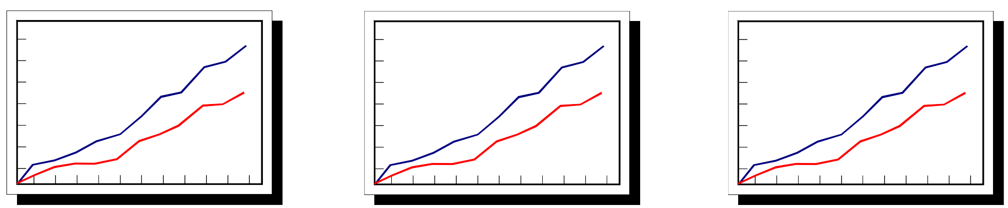
1. Split the dataset randomly into halves

Master Dataset

Analysis Dataset

Test Dataset

Diagnostics



Put half of the claimants into an analysis dataset and half into a test dataset. This is to prevent over-fitting. The scoring will be constructed on the analysis dataset and tested on the test dataset. Diagnostic reports are run on each dataset and compared to each other to ensure that the compositions of the datasets are essentially similar. Reports are run on age, sex, cost, as well as disease and Rx markers.

SCIOinspire FRESH thinking REAL results

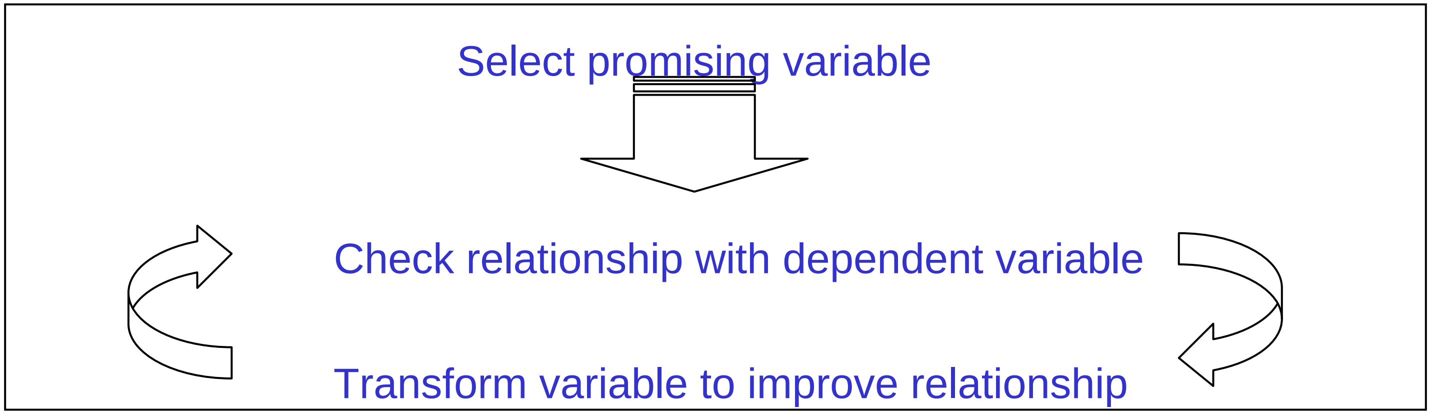
2. Build and Transform independent variables

- In any data-mining project, the output is only as good as the input.
- Most of the time and resources in a data mining project are actually used for variable preparation and evaluation, rather than generation of the actual “recipe”.

3. Build composite dependent variable

- A key step is the choice of dependent variable.   What is the best choice?
- A likely candidate is total patient cost in the predictive period. But total cost has disadvantages
  - It includes costs such as injury or maternity that are not generally predictable.
  - It includes costs that are steady and predictable, independent of health status (capitated expenses).
  - It may be affected by plan design or contracts.
- We generally predict total cost (allowed charges) net of random costs and capitated expenses.
- Predicted cost can be converted to a risk-factor.

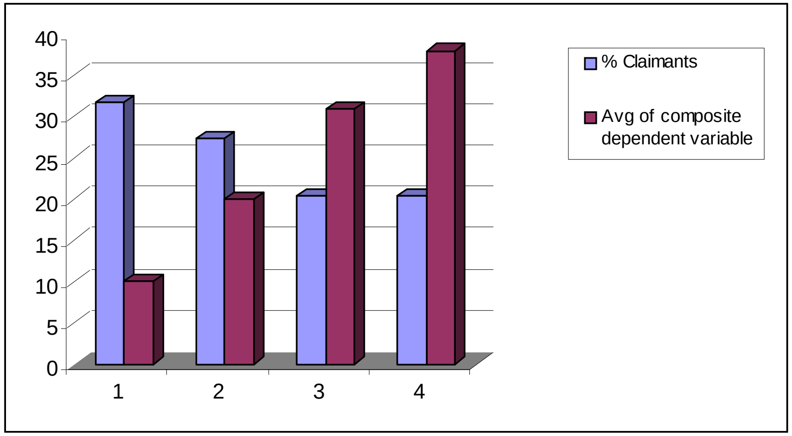
3. Build and transform Independent Variables



- Typical transforms include
- Truncating data ranges to minimized the effects of outliers.
- Converting values into binary flag variables.
- Altering the shape of the distribution with a log transform to compare orders of magnitude.
- Smoothing progression of independent variables

3. Build and transform Independent Variables

- A simple way to look at variables
- Convert to a discrete variable. Some variables such as number of prescriptions are already discrete. Real-valued variables, such as cost variables, can be grouped into ranges
- Each value or range should have a significant portion of the patients.
- Values or ranges should have an ascending or descending relationship with average value of the composite dependent variable.



Typical "transformed" variable

4. Select Independent Variables

- The following variables were most promising
- Age -Truncated at 15 and 80
- Baseline cost
- Number of comorbid condition truncated at 5
- MClass
  - Medical claims-only generalization of the comorbidity variable.
  - Composite variable that counts the number of distinct ICD9 ranges for which the claimant has medical claims.
  - Ranges are defined to separate general disease/condition categories.
- Number of prescriptions truncated at 10

4. Select Independent Variables (contd.)

- Scheduled drug prescriptions truncated at 5
- NClass
  - Rx-only generalization of the comorbidity variable.
  - Composite variable that counts the number of distinct categories distinct ICD9 ranges for which the claimant has claims.
  - Ranges are defined using GPI codes to separate general disease/condition categories.
- Ace inhibitor flag
- Anticoagulants flag
- Diuretics flag
- Number of corticosteroid drug prescriptions truncated at 2

Neuroleptic drug flag

Digoxin flag

5. Run Stepwise Linear Regression

An ordinary linear regression is simply a formula for determining a best-possible linear equation describing a dependent variable as a function of the independent variables. But this pre-supposes the selection of a best-possible set of independent variables. How is this best-possible set of independent variables chosen?

One method is a stepwise regression. This is an algorithm that determines both a set of variables and a regression. Variables are selected in order according to their contribution to incremental R^2

5. Run Stepwise Linear Regression (continued)

Stepwise Algorithm

1. Run a single-variable regression for each independent variable. Select the variable that results in the greatest value of R^2. This is "Variable 1".

2. Run a two-variable regression for each remaining independent variable. In each regression, the other independent variable is Variable 1. Select the remaining variable that results in the greatest <u>incremental</u> value of R^2. This is "Variable 2."

3. Run a three-variable regression for each remaining independent variable. In each regression, the other two independent variables are Variables 1 and 2. Select the remaining variable that results in the greatest <u>incremental</u> value of R^2. This is "Variable 3."

……

n. Stop the process when the incremental value of R^2 is below some pre-defined threshold.

6. Results - Examples

- Stepwise linear regressions were run using the "promising" independent variables as inputs and the composite dependent variable as an output.
- Separate regressions were run for each patient sex.
- Sample Regressions

- Female

Scheduled drug prescription	358.1
NClass	414.5
MClass	157.5
Baseline cost	0.5
Diabetes Dx	1818.9
Intercept	18.5

Why are some variables selected while others are omitted? The stepwise algorithm favors variables that are relatively uncorrelated with previously-selected variables. The variables in the selections here are all relatively independent of each other.

6. Results - Examples

- Examples of application of the female model

Female Regression Regression Formula

(Scheduled Drug *358.1) + (NClass*414.5) + (Cost*0.5) + (Diabetes*1818.9) + (MClass*157.5) -18.5

Claimant ID	Raw Value	Transformed Value	Predicted Value	Actual Value
	Schedule Drugs			
1	3	2	$ 716.20	
2	2	2	$ 716.20	
3	0	1	$ 358.10	
	NClass			
1	3	3	$ 1,243.50	
2	6	6	$ 2,487.00	
3	0	0.5	$ 207.25	
	Cost			
1	423	2,000	$ 1,000.00	
2	5,244	6,000	$ 3,000.00	
3	1,854	2,000	$ 1,000.00	
	Diabetes			
1	0	0	$ -	
2	0	0	$ -	
3	0	0	$ -	
	MClass			
1	8	3	$ 472.50	
2	3	2	$ 315.00	
3	0	0.5	$ 78.75	
	TOTAL			
1			$ 3,413.70	$ 4,026.00
2			$ 6,499.70	$ 5,243.00
3			$ 1,625.60	$ 1,053.00

Transform Function

	Schedule Drugs		
Value Range	RV< 2	2 < RV < 5	RV >5
Transformed Value	1.0	2.0	3.0

	NClass		
Value Range	RV < 2	2 < RV < 5	RV > 5
Transformed Value	0.5	3.0	6.0

	Cost		
Value Range	RV < 5k	;k < RV < 10l	RV > 10k
Transformed Value	2,000	6,000	10,000

	Diabetes	
Value Range	Yes	No
Transformed Value	1.0	0.0

	MClass		
Value Range	RV < 1	1 < RV < 7	RV > 7
Transformed Value	0.5	2.0	3.0

Evaluation – Case Examples

Background - Case 1

- Large client.
- Several years of data provided for modeling.
- Never able to become comfortable with data which did not perform well according to our benchmark statistics ($/claimant; $pmpm; number of claims per member).

BENCHMARK DATA		(Commercial only)	pmpm	Claims/ member/ year
		Medical Only	$ 70.40	14.40
		Rx Only	$ 16.49	7.70
		TOTAL	$ 86.89	22.10

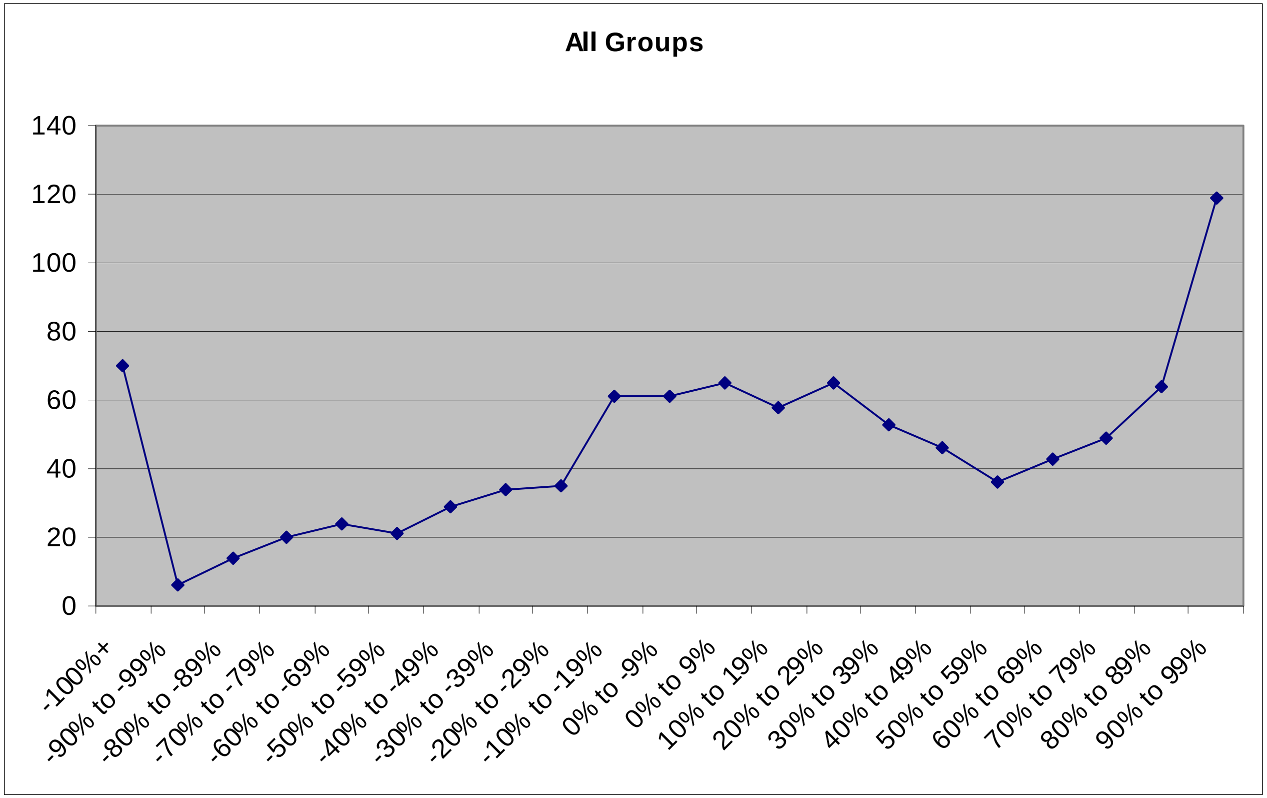
CLIENT DATA		(Commercial; excludes Capitation)	pmpm	Claims/ member/ year
		Medical + Rx	$ 32.95	5.36
		TOTAL	$ 32.95	5.36

SCIOinspire
FRESH thinking REAL results

Background – Case 1

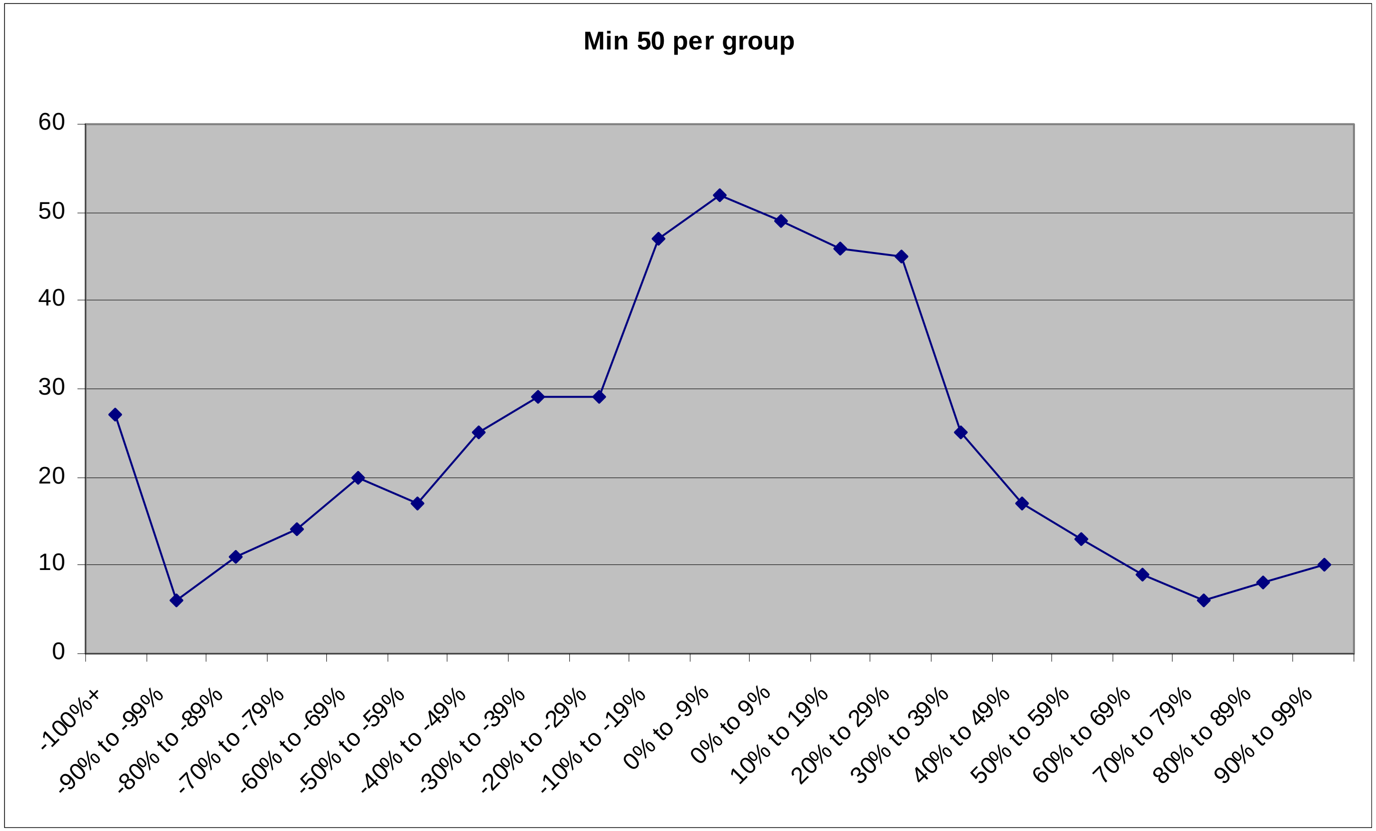
- Built models to predict cost in year 2 from year 1.
- Now for the hard part: evaluating the results.

How well does the model perform?



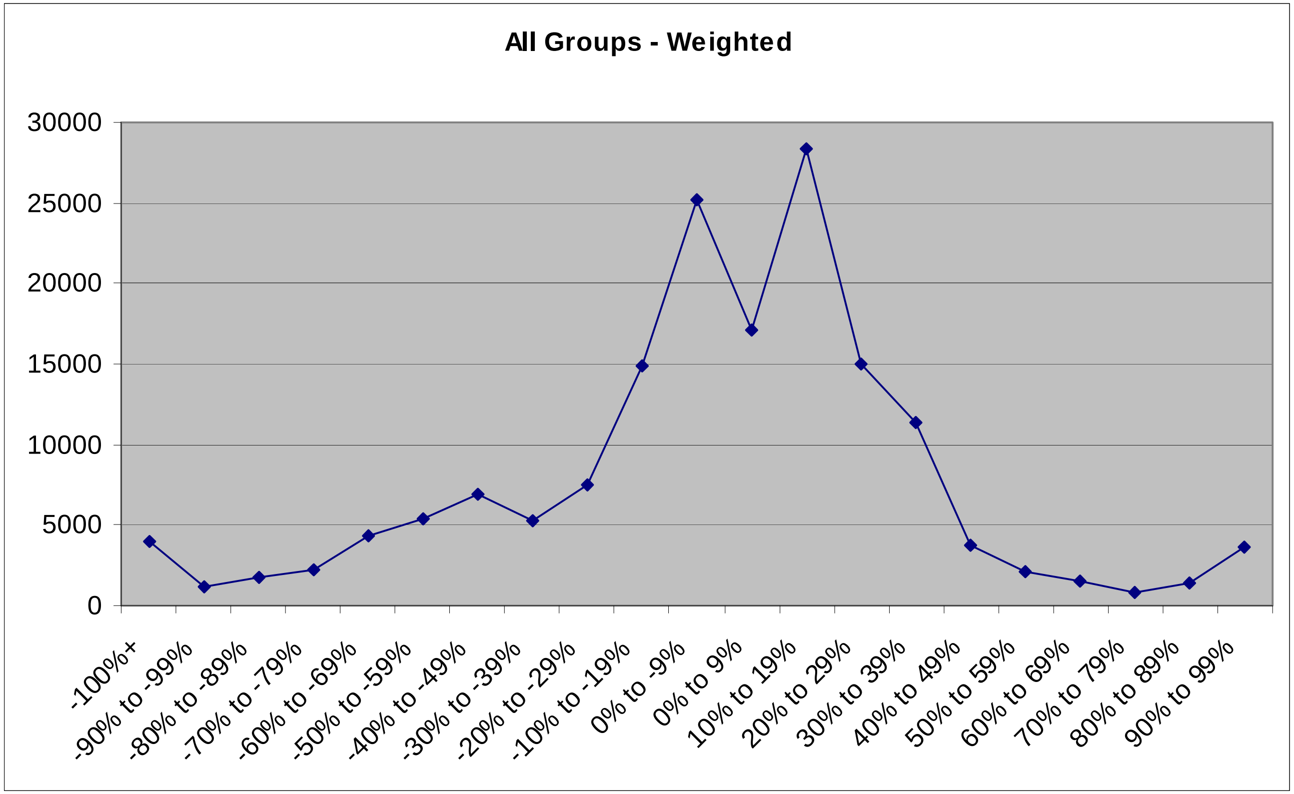
Analysis 1: all groups. This analysis shows that, at the group level, prediction is not particularly accurate, with a significant number of groups at the extremes of the distribution.

How well does the model perform?



Analysis 2: Omitting small groups (under 50 lives) significantly improves the actual/predicted outcomes.

How well does the model perform?



Analysis 3: Weighting the results by the number of lives in the group shows that most predictions lie within +/- 30% of the actual.

Conclusion

- Significant data issues were identified and not resolved.
- This was a large group carrier who had many groups "re-classified" during the period. They were unable to provide good data that "matched" re-classified groups to their previous numbers.
- Conclusion: if you are going to do anything in this area, be sure you have good data.

SCIOinspire
FRESH thinking REAL results

Background – Case 2.

- Client uses a manual rate basis for rating small cases. Client believes that case selection/ assignment may result in case assignment to rating classes that is not optimal.

- A predictive model may add further accuracy to the class assignment process and enable more accurate rating and underwriting to be done.

Background

- A number of different tree models were built (at client's request).
- Technically, an optimal model was chosen.

Problem: how to convince Underwriting that:

- Adding the predictive model to the underwriting process produces more accurate results; and
- They need to change their processes to incorporate the predictive model.

Some data

Node	PREDICTED Average Profit	PREDICTED Number in Node	PREDICTED Number in Node (Adjusted)	ACTUAL Number in node	ACTUAL Average Profit
1	(3.03)	70	173	170	(0.60)
2	0.19	860	2,122	2,430	0.07
3	(0.20)	2,080	5,131	6,090	(0.06)
4	0.09	910	2,245	2,580	0.10
5	(0.40)	680	1,678	20	0.02
6	(0.27)	350	863	760	0.16
7	0.11	650	1,604	1,810	0.04
8	0.53	190	469	470	(0.01)
9	(0.13)	1,150	2,837	2,910	0.03
10	0.27	1,360	3,355	3,740	0.04
11	0.38	1,560	3,849	3,920	(0.07)
12	0.08	320	789	830	0.08
13	0.06	12,250	30,221	29,520	0.02
14	0.27	2,400	5,921	6,410	0.21
15	(1.07)	540	1,332	1,320	(0.03)
16	0.07	10,070	24,843	24,950	(0.08)
17	(0.33)	1,400	3,454	3,250	(0.10)
18	0.11	4,460	11,003	11,100	0.08
19	(0.13)	1,010	2,492	2,100	(0.11)
		42,310	104,380	104,380	0.005

SCIOinspire FRESH thinking REAL results

How well does the model perform?

Node	PREDICTED Average Profit	PREDICTED Number in Node	PREDICTED Number in Node (Adjusted)	ACTUAL Number in node	ACTUAL Average Profit	Directionally Correct (+ or -)
1	(3.03)	70	173	170	(0.60)	
2	0.19	860	2,122	2,430	0.07	
3	(0.20)	2,080	5,131	6,090	(0.06)	
4	0.09	910	2,245	2,580	0.10	
5	(0.40)	680	1,678	20	0.02	
6	(0.27)	350	863	760	0.16	
7	0.11	650	1,604	1,810	0.04	
8	0.53	190	469	470	(0.01)	
9	(0.13)	1,150	2,837	2,910	0.03	
10	0.27	1,360	3,355	3,740	0.04	
11	0.38	1,560	3,849	3,920	(0.07)	
12	0.08	320	789	830	0.08	
13	0.06	12,250	30,221	29,520	0.02	
14	0.27	2,400	5,921	6,410	0.21	
15	(1.07)	540	1,332	1,320	(0.03)	
16	0.07	10,070	24,843	24,950	(0.08)	
17	(0.33)	1,400	3,454	3,250	(0.10)	
18	0.11	4,460	11,003	11,100	0.08	
19	(0.13)	1,010	2,492	2,100	(0.11)	
		42,310	104,380	104,380	0.005	

6 red
13 green

How well does the model perform?

Node	PREDICTED Average Profit	PREDICTED Number in Node	PREDICTED Number in Node (Adjusted)	ACTUAL Number in node	ACTUAL Average Profit	Directionally Correct (+ or -)	Predicted to be Profitable
1	(3.03)	70	173	170	(0.60)		
2	0.19	860	2,122	2,430	0.07		
3	(0.20)	2,080	5,131	6,090	(0.06)		
4	0.09	910	2,245	2,580	0.10		
5	(0.40)	680	1,678	20	0.02		
6	(0.27)	350	863	760	0.16		
7	0.11	650	1,604	1,810	0.04		
8	0.53	190	469	470	(0.01)		
9	(0.13)	1,150	2,837	2,910	0.03		
10	0.27	1,360	3,355	3,740	0.04		
11	0.38	1,560	3,849	3,920	(0.07)		
12	0.08	320	789	830	0.08		
13	0.06	12,250	30,221	29,520	0.02		
14	0.27	2,400	5,921	6,410	0.21		
15	(1.07)	540	1,332	1,320	(0.03)		
16	0.07	10,070	24,843	24,950	(0.08)		
17	(0.33)	1,400	3,454	3,250	(0.10)		
18	0.11	4,460	11,003	11,100	0.08		
19	(0.13)	1,010	2,492	2,100	(0.11)		
		42,310	104,380	104,380	0.005	6 red 13 green	11 nodes

Underwriting Decision-making

Underwriting Decision	Total Profit	Average Profit per Case	Cases Written
Accept all cases as rated.	557.5	0.005	104,380

Underwriting Decision-making

Underwriting Decision	Total Profit	Average Profit per Case	Cases Written
Accept all cases as rated.	557.5	0.005	104,380
Accept all cases predicted to be profitable; reject all predicted unprofitable cases.	1,379.4	0.016	87,760

SCIOinspire
FRESH thinking REAL results

Underwriting Decision-making

Underwriting Decision	Total Profit	Average Profit per Case	Cases Written
Accept all cases as rated.	557.5	0.005	104,380
Accept all cases predicted to be profitable; reject all predicted unprofitable cases.	1,379.4	0.016	87,760
Accept all cases predicted to be profitable; rate all cases predicted to be unprofitable +10%.	2,219.5	0.021	104,380

SCIOinspire
FRESH thinking REAL results

Underwriting Decision-making

Underwriting Decision	Total Profit	Average Profit per Case	Cases Written
Accept all cases as rated.	557.5	0.005	104,380
Accept all cases predicted to be profitable; reject all predicted unprofitable cases.	1,379.4	0.016	87,760
Accept all cases predicted to be profitable; rate all cases predicted to be unprofitable +10%.	2,219.5	0.021	104,380
Accept all cases for which the directional prediction is correct.	2,543.5	0.026	100,620

SCIOinspire FRESH thinking REAL results

Underwriting Decision-making

Underwriting Decision	Total Profit	Average Profit per Case	Cases Written
Accept all cases as rated.	557.5	0.005	104,380
Accept all cases predicted to be profitable; reject all predicted unprofitable cases.	1,379.4	0.016	87,760
Accept all cases predicted to be profitable; rate all cases predicted to be unprofitable +10%.	2,219.5	0.021	104,380
Accept all cases for which the directional prediction is correct.	2,543.5	0.026	100,620
Accept all cases for which the directional prediction is correct; rate predicted unprofitable cases by +10%	3,836.5	0.038	100,620

SCIOinspire FRESH thinking REAL results

Underwriting Decision-making

Underwriting Decision	Total Profit	Average Profit per Case	Cases Written
Accept all cases as rated.	557.5	0.005	104,380
Accept all cases predicted to be profitable; reject all predicted unprofitable cases.	1,379.4	0.016	87,760
Accept all cases predicted to be profitable; rate all cases predicted to be unprofitable +10%.	2,219.5	0.021	104,380
Accept all cases for which the directional prediction is correct.	2,543.5	0.026	100,620
Accept all cases for which the directional prediction is correct; rate predicted unprofitable cases by +10%	3,836.5	0.038	100,620
Accept all cases for which the directional prediction is correct.	2,540.8	0.025	101,090

SCIOinspire FRESH thinking REAL results

Example 3: evaluating a high-risk model

Background

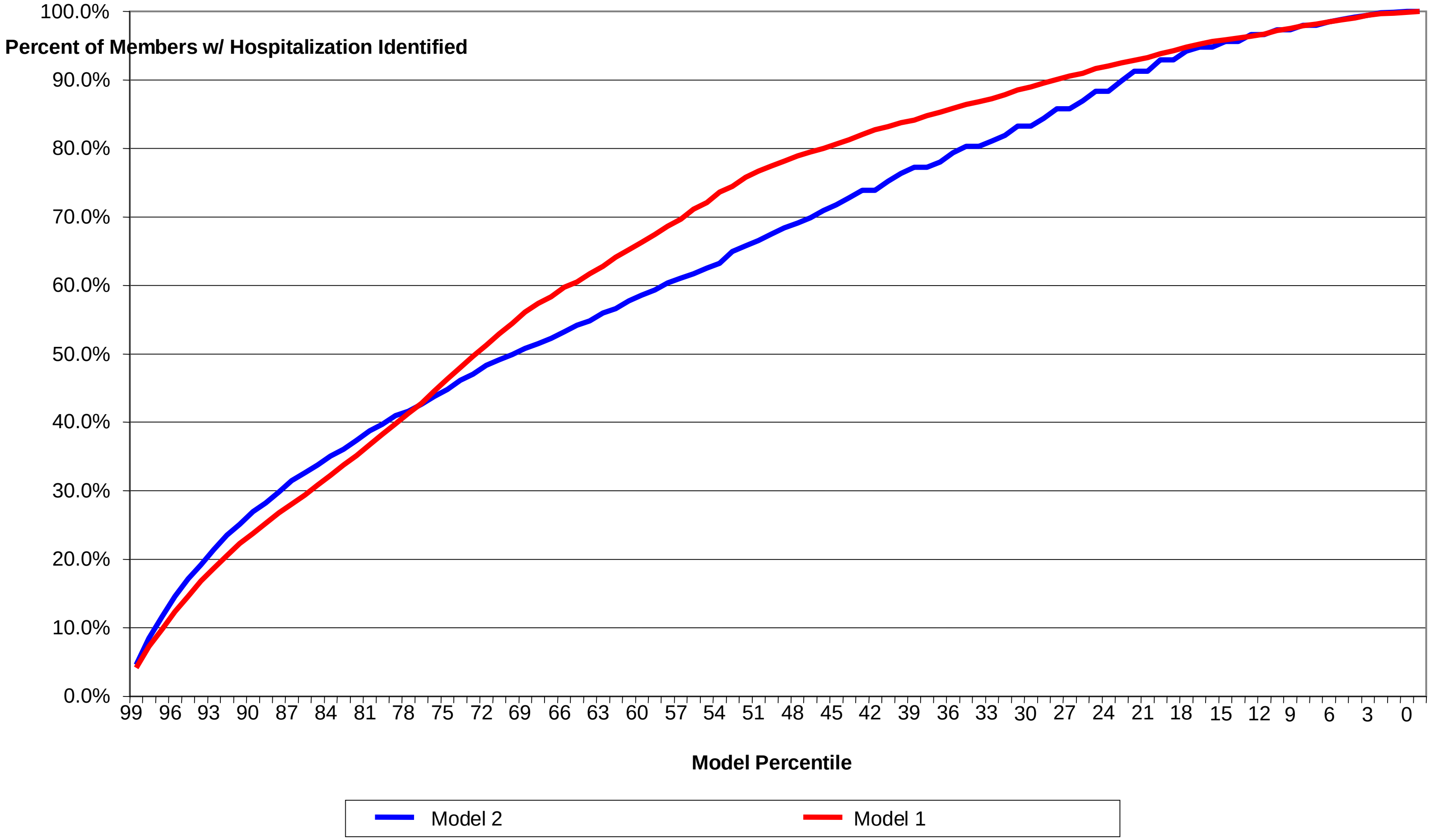
- Large health plan client seeking a model to improve case identification for case management.
- Considered two commercially-available models:
  - Version 1: vendor's typical predictive model based on conditions only. Model is more typically used for risk-adjustment (producing equivalent populations).
  - Version 2: vendor's high-risk predictive model that predicts the probability of a member having an event in the next 6-12 months.

SCIOinspire
FRESH thinking REAL results

Analysis

- Client initially rejected model 2 as not adding sufficient value compared with model 1. (Vendor's pricing strategy was to charge additional fees for model 2) based on cumulative predictions.

Lift Chart – Comparison between Two models

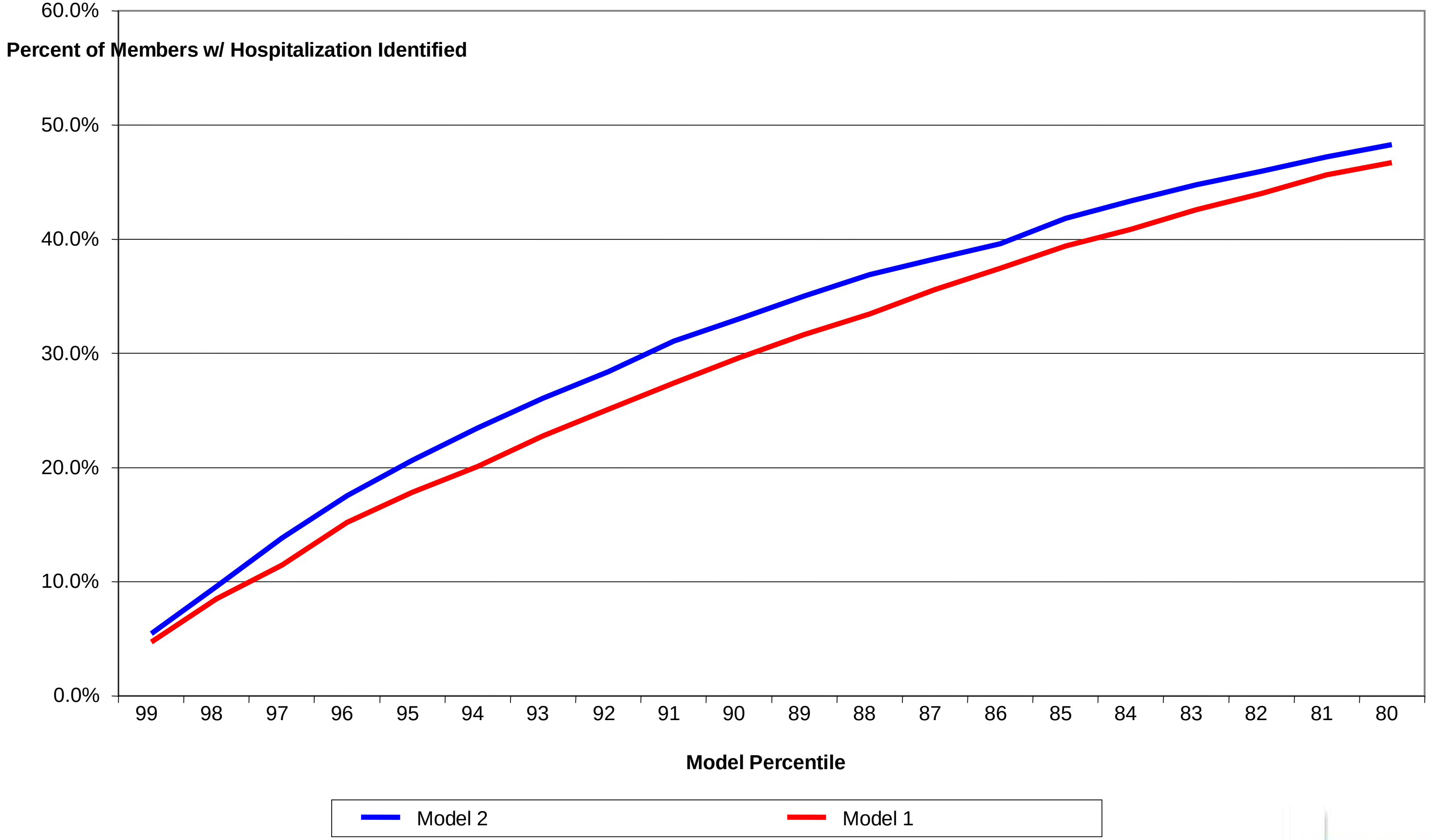


Analysis

- Looked at over a narrower range, however, the results appear different.

Background

Lift Chart – Comparison between Two models



Analysis

Decile		Decile Admissions					
From	To	Population	Expected	Actual	Predicted Frequency	Actual Frequency	Predictive ratio
100%	90%	1,690	808	694	47.8%	41.1%	85.9%
90%	80%	1,699	268	321	15.8%	18.9%	119.6%
80%	70%	1,657	152	247	9.2%	14.9%	162.0%
70%	60%	1,673	107	191	6.4%	11.4%	178.4%
60%	50%	1,681	82	168	4.9%	10.0%	204.0%
50%	40%	1,760	67	165	3.8%	9.4%	246.7%
40%	30%	1,667	50	118	3.0%	7.1%	236.0%
30%	20%	1,729	38	92	2.2%	5.3%	241.9%
20%	10%	1,624	26	68	1.6%	4.2%	261.7%
10%	0%	1,708	91	37	5.3%	2.2%	40.9%
		16,888	1,690	2,101	100%	124.4%	

SCIOinspire FRESH thinking REAL results

Example 4: a wellness model

Solucia Wellness Model

- Using data from a large health plan (multi-million lives; both self-reported data and health claims) we developed a risk-factor model that relates claims dollars to risk factors;
- Multiple regression model;
- 15 different risk factors;
- Multiple categorical responses.

Solucia Wellness Model

Attribute	Variable	Values	Cost Impact
	Intercept	1	190
Personal Disease History 1	Chronic Obstructive Pulmonary Disease (COPD), Congestive Heart Failure (CHF), Coronary Heart Disease (CHD), Peripheral Vascular Disease (PVD) and Stroke	0 (No)	-
		1 (Yes)	10,553
Health Screenings	Have you had a SIGMOIDOSCOPY within the last 5 years? (tube inserted in rectum to check for lower intestine problems)	0 (No)	-
		1 (Yes)	2,045
Weight Management	Body Mass Index	26 (Min)	3,069
		40 (No Value)	4,722
		45 (Max)	5,312
Health Screenings	Influenza (flu) within the last 12 months?	0 (No)	-
		1 (Yes)	1,176
Personal Disease History 2	Have you never been diagnosed with any of the following: list of 27 major conditions	0 (No)	-
		1 (Yes)	(1,220)
Personal Disease History 3	TIA (mini-stroke lasting less than 24 hrs), Heart Attack, Angina, Breast Cancer, Emphysema	0 (No)	-
		1 (Yes)	2,589
Immunizations	Pneumonia	0 (No)	-
		1 (Yes)	1,118
Physical Activity 1	Moderate-intensity physical activity - minutes per day	0 (Min, No Value)	-
		20 (Max)	(915)
Stress and Well-Being	In the last month, how often have you been angered because of things that happened that were outside your control?	0 (Never, Almost Never, Sometimes, Fairly Often)	-
		1 (Very Often, No Value)	1,632

Solucia Wellness Model

Skin Protection	Please rate how confident you are that you can have your skin checked by a doctor once a year?	1 (Not at all confident)	(224)
		2 (Not confident)	(447)
		3 (Fairly confident)	(671)
		4 (Confident)	(894)
		5 (Very Confident)	(1,118)
		7 (No Value)	(1,565)
Women's health 1	Are you currently on hormone replacement therapy (Estrogen Therapy, Premarin) or planning to start?	0 (No)	-
		1 (Yes)	999
Women's health 2	Select the appropriate answer regarding pregnancy status/plan	1 (NotPlanning (I am planning on becoming pregnant in the next 6 months.))	590
		2 (No Value)	1,181
		3 (Planning (I am planning on becoming pregnant in the next 6 months.))	1,771
		4 (Pregnant (I am currently pregnant))	2,361
Physical Activity 2	HIGH intensity activities? (hours per week)	0 (Min, No Value)	-
		3 (Max)	(917)
Nutrition	On a typical day, how many servings do you eat of whole grain or enriched bread, cereal, rice, and pasta?	0 (None, No Value)	-
		1 (OneThree, FourFive)	(868)
		2 (SixPlus)	(1,736)
Tobacco	Please rate how confident you are that you can keep from smoking cigarettes when you feel you need a lift.	1 (Not at all confident)	(294)
		1.5 (No Value)	(441)
		2 (Not confident)	(588)
		3 (Fairly confident)	(883)
		4 (Confident)	(1,177)

Discussion?

Selected references

This is not an exhaustive bibliography. It is only a starting point for explorations.

- Shapiro, A.F. and Jain, L.C. (editors); *Intelligent and Other Computational Techniques in Insurance;* World Scientific Publishing Company; 2003.
- Dove, Henry G., Duncan, Ian, and Robb, Arthur; *A Prediction Model for Targeting Low-Cost, High-Risk Members of Managed Care Organizations*; The American Journal of Managed Care, Vol 9 No 5, 2003
- Berry, Michael J. A. and Linoff, Gordon; *Data Mining Techniques for Marketing, Sales and Customer Support*; John Wiley and Sons, Inc; 2004
- Montgomery, Douglas C., Peck, Elizabeth A., and Vining, G Geoffrey; *Introduction to Linear Regression Analysis*; John Wiley and Sons, Inc; 2001
- Kahneman, Daniel, Slovic, Paul, and Tversky (editors); *Judgment under uncertainty: Heuristics and Biases;* Cambridge University Press; 1982

Selected references (contd.)

- Dove, Henry G., Duncan, Ian, and others; *Evaluating the Results of Care Management Interventions: Comparative Analysis of Different Outcomes Measures.* The SOA study of DM evaluation, available on the web-site at

  http://www.soa.org/professional-interests/health/hlth-evaluating-the-results-of-care-management-interventions-comparative-analysis-of-different-outcomes-measures-claims.aspx

- Winkelman R. and S. Ahmed. *A comparative analysis of Claims Based Methods of health risk assessment ofr Commercial Populations*. (2007 update to the SOA Risk-Adjuster study.) Available from the SOA; the 2002 study is on the website at: http://www.soa.org/files/pdf/_asset_id=2583046.pdf.

Further Questions?

iduncan@soluciaconsulting.com

Solucia Inc.

220 Farmington Avenue, Suite 4

Farmington, CT 06032

860-676-8808

www.soluciaconsulting.com