SOLUCIA, INC.

Introduction to Predictive Modeling

June 19, 2008

1

Introductions

SOLUCIA, INC.

Ian Duncan, FSA, FIA, FCIA, MAAA.

President, Solucia Consulting, a SCIOinspire company.

Farmington, CT.

2

Introduction / Objective

SOLUCIA, INC.

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

SOLUCIA, INC.

Predictive Modeling:
A Review of the Basics

Definition of Predictive Modeling

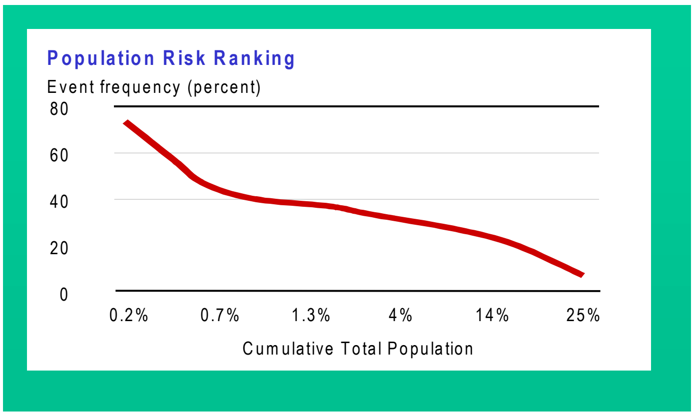
SOLUCIA, INC.

"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"

SOLUCIA, INC.



PM – more often wrong than right…

SOLUCIA, INC.

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

-- *Harvard Economic Service, December 1929*

7

Why do it? Potential Use of Models

SOLUCIA, INC.

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.**

8

Identification – how?

SOLUCIA, INC.

- The art and science of predictive modeling!
- 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

SOLUCIA, INC.

Codes are required for payment, so they tend to be accurate. The 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); etc. etc.

They are the "raw material" of predictive modeling.

"Grouper" models take the raw material and consolidate it into manageable like categories.

Identification – example (Diabetes)

SOLUCIA, INC.

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.

Diabetes – additional codes

SOLUCIA, INC.

CODES DIABETES;	CODE TYPE	DESCRIPTION - ADDITIONAL
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

SOLUCIA, INC.

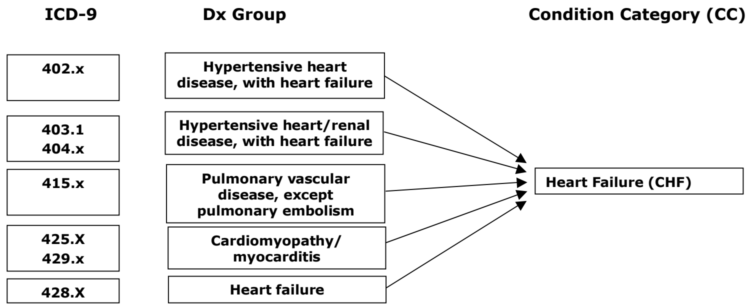
Insulin or Oral Hypoglycemic Agents are often used to identify members. A simple example follows; for more detail, see the HEDIS code-set.

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**

Grouper Models

SOLUCIA, INC.



- 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;
- Scores can range from 0.3 (for young people without diagnoses) to numbers in the 40’s and 50’s (for multiple co-morbid patients).

Construction of a model

SOLUCIA, INC.

Finding	Hierarchy	Coefficient	Notes
Diabetes	Low-cost Diabetes	0.000	Trumped by High-cost Diabetes
Diabetic Nephropathy		2.455	
Angina	Low-cost CAD	0.000	Trumped by High-cost CAD
CAD	CAD	3.500	
Migraine	Med. Cost headache	0.208	
	SUBTOTAL	6.163	
Age-related Base		0.306	
Gender-related Base		-0.087	
	RISK SCORE	6.382	

All people are not equally identifiable

SOLUCIA, INC.

Definition Examples:

Narrow: **Hospital Inpatient (primary Dx); Face-to-face professional (no X-Ray; Lab)**

Broad: **Hospital I/P (any Dx); All professional**

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%

Identification: False Positives/ False Negatives

SOLUCIA, INC.

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

	Year 1	Narrow	+ Broad	+ 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%

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%

Why do it? Potential Use of Models

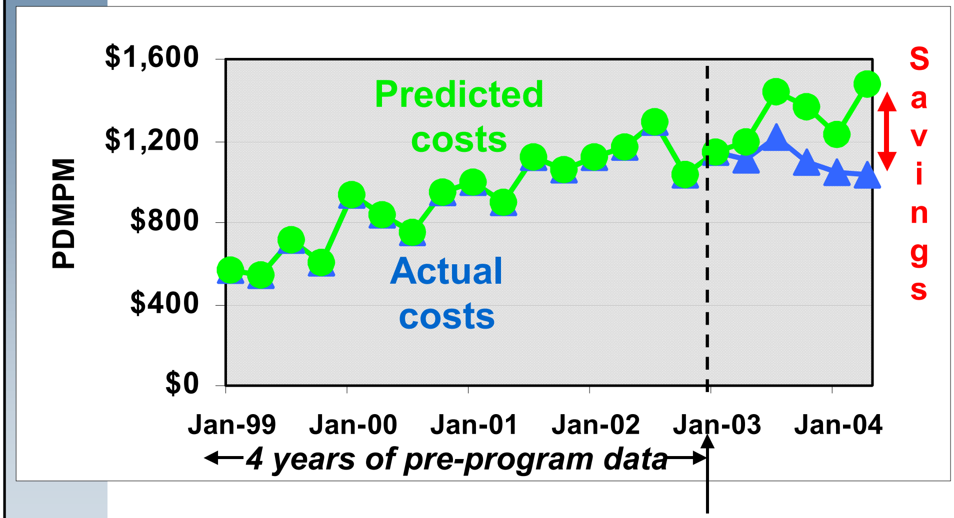
SOLUCIA, INC.

Program Evaluation/ Reimbursement Perspective

- **Predicting *what would have happened* absent a program.**
- **Predicting resource use in the "typical" population.**

19

Example 1: Time Series



20

Example 2: 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

21

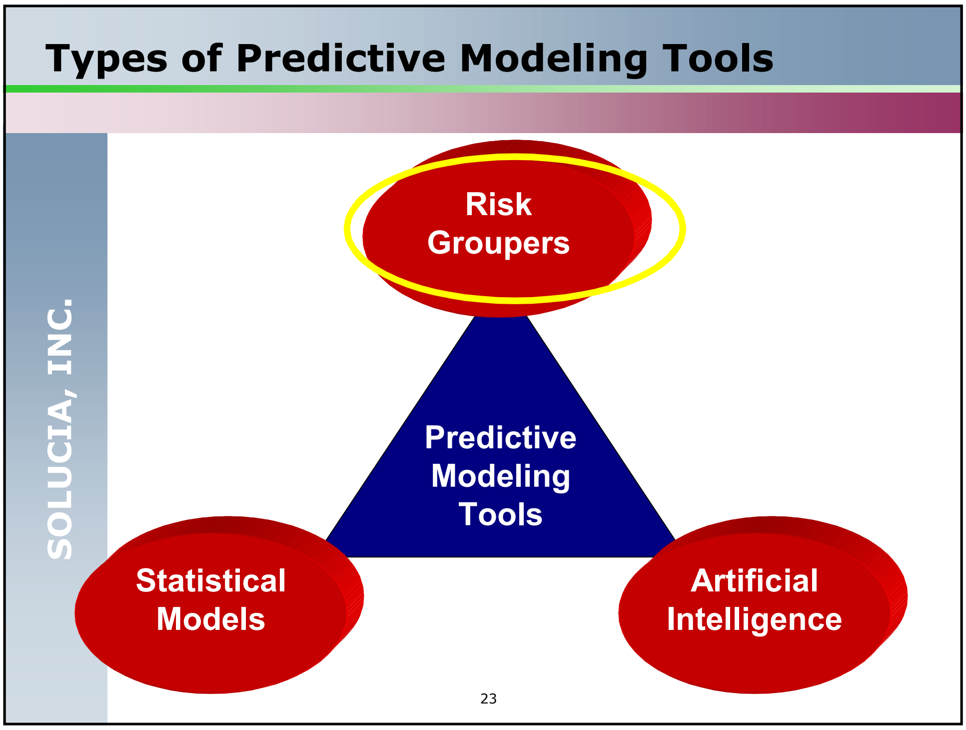
Why do it? Potential Uses of Models

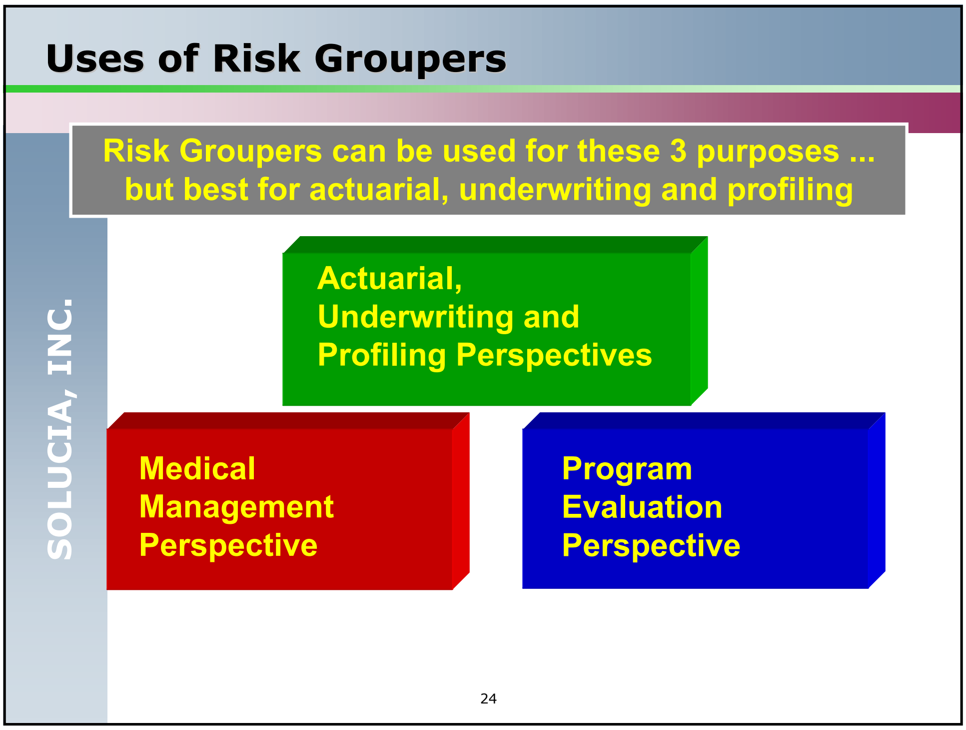
SOLUCIA, INC.

Actuarial, Underwriting and Profiling Perspectives

- Calculating renewal premium
- Profiling of provider
- Provider & health plan contracting

22


Types of Predictive Modeling Tools
SOLUCIA, INC.
Risk Groupers
Predictive Modeling Tools
Statistical Models
Artificial Intelligence
23


Uses of Risk Groupers
SOLUCIA, INC.
Risk Groupers can be used for these 3 purposes ... but best for actuarial, underwriting and profiling
Actuarial, Underwriting and Profiling Perspectives
Medical Management Perspective
Program Evaluation Perspective
24

Risk Groupers

SOLUCIA, INC.

What are the different types of risk groupers?

Selected Risk Groupers

SOLUCIA, INC.

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

Risk Grouper Summary

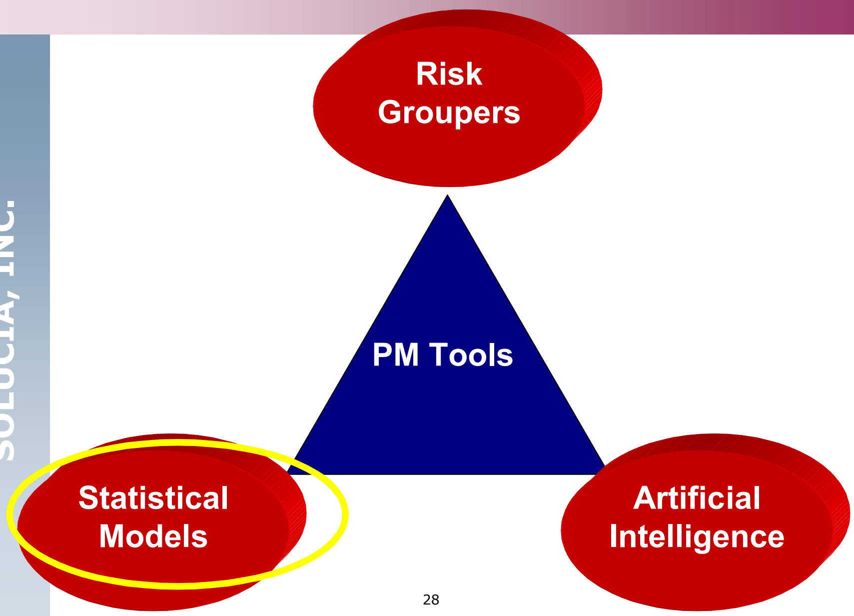
SOLUCIA, INC.

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

SOLUCIA, INC.



Uses of Statistical Models

Statistical models can be used for all 3 uses

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

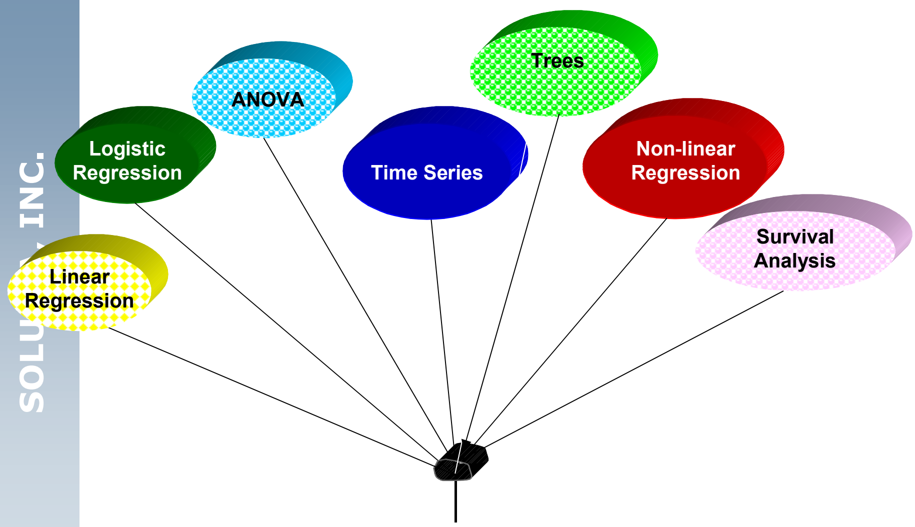
SOLUC

Statistical Models

SOLUCIA, INC.

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

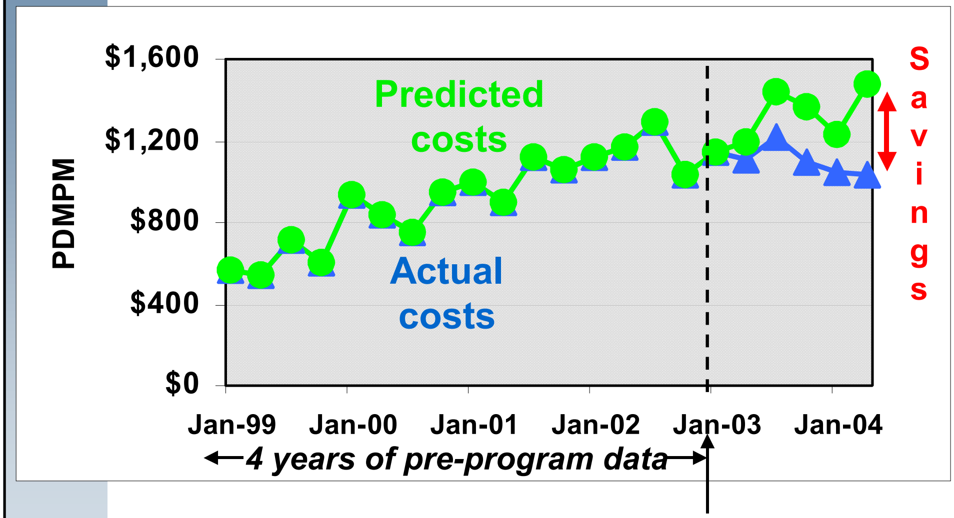
SOLUCIA, INC.

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



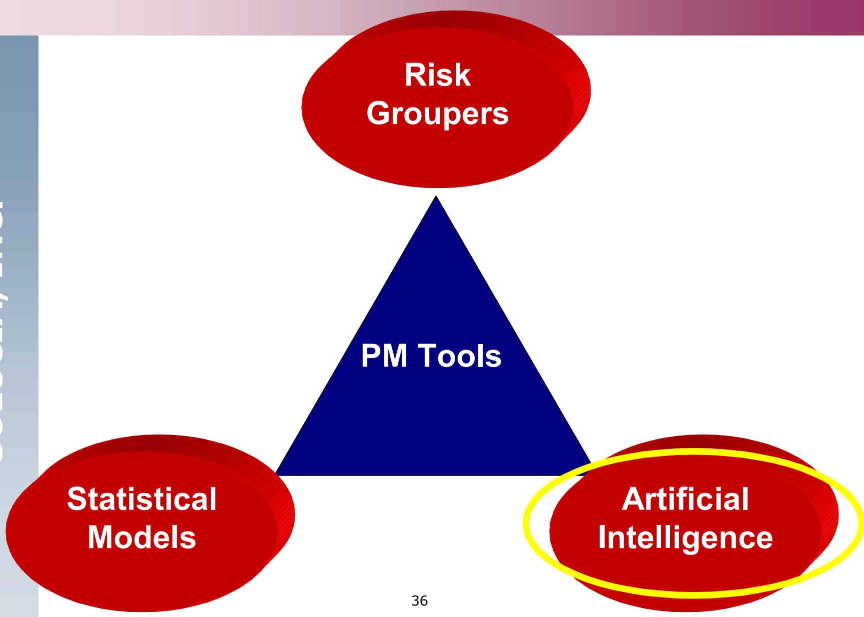
Statistical Model Summary

SOLUCIA, INC.

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).

Types of Predictive Modeling Tools

SOLUCIA, INC.



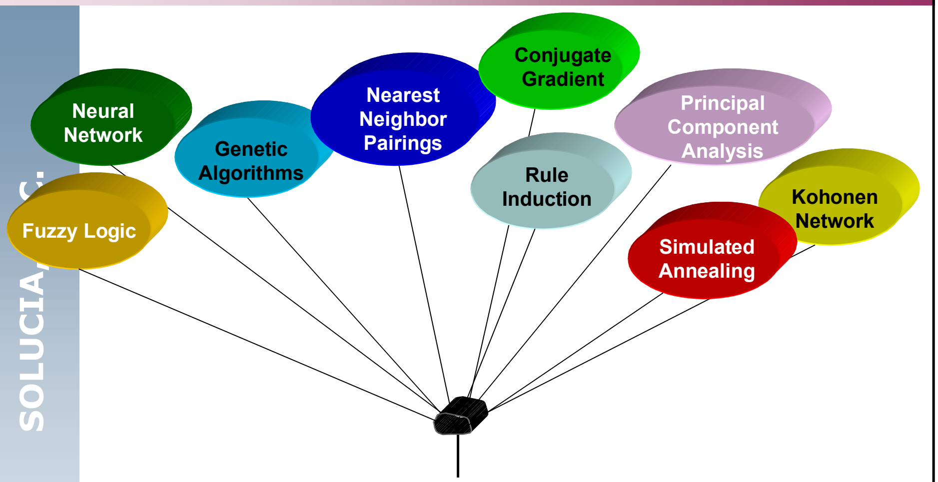
Artificial Intelligence Models

SOLUCIA, INC.

What are the different types of artificial intelligence models?

37

Artificial Intelligence Models



38

Features of Neural Networks

SOLUCIA, INC.

Reality

NN tracks complex relationships by resembling the human brain

Perception

NN can accurately model complicated health care systems

Reality

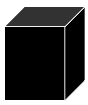
- Performance equals standard statistical models
- Models overfit data

39

Neural Network Summary

SOLUCIA, INC.

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).



40

In Summary

SOLUCIA, INC.

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. impactability).
4. "Mixed" models can increase the power of a single model.

41

PM is NOT always about *Cost Prediction…..*

SOLUCIA, INC.

…..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?

42

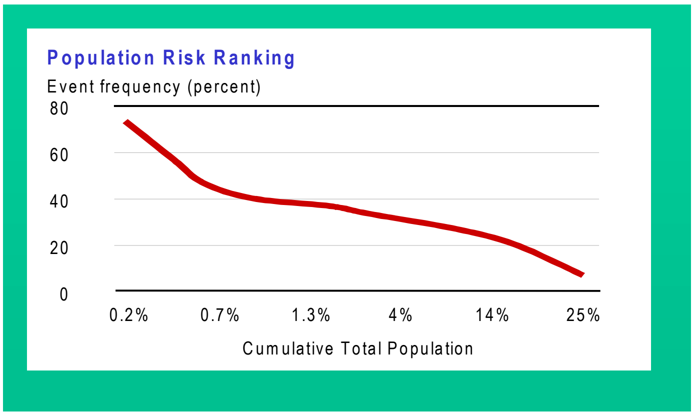
Remember this chart?

	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%

43

Decreasing Cost / Decreasing Opportunity

SOLUCIA, INC.



44

Economic Model: Simple example

SOLUCIA, INC.

- 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.

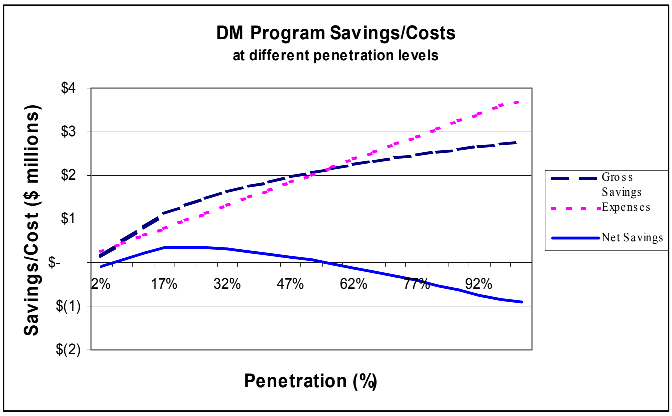
Key drivers of the economic model

SOLUCIA, INC.

- 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

SOLUCIA, INC.



SOLUCIA, INC.

Modeling

What is a model?

SOLUCIA, INC.

- 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.

SOLUCIA, INC.

Evaluation – Case Examples

Background – Case 1

SOLUCIA, INC.

- 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

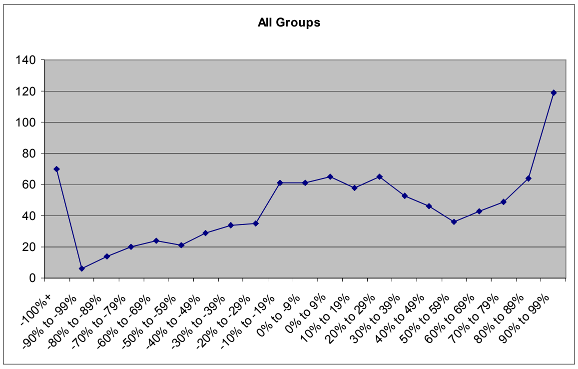
Background – Case 1

SOLUCIA, INC.

- 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?

SOLUCIA, INC.

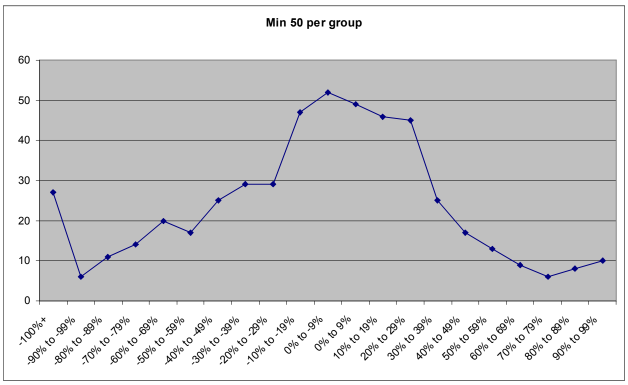


Analysis 1: all groups. This analysis shows that, at the group level, prediction is not particularly accurate, with a significant number of groups at the extremes of the distribution.

53

How well does the model perform?

SOLUCIA, INC.

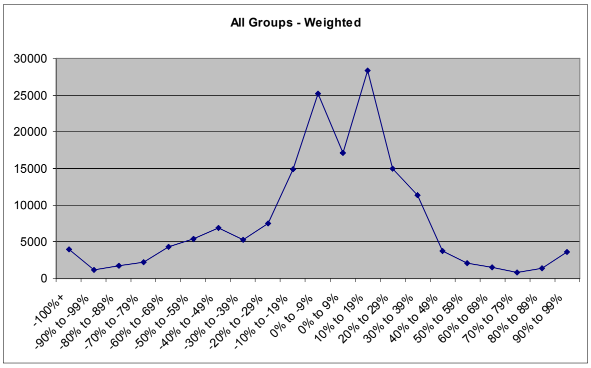


Analysis 2: Omitting small groups (under 50 lives) significantly improves the actual/predicted outcomes.

54

How well does the model perform?

SOLUCIA, INC.



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

Conclusion

SOLUCIA, INC.

- 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.

Background – Case 2.

SOLUCIA, INC.

- 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

SOLUCIA, INC.

- 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

SOLUCIA, INC.

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

How well does the model perform?

SOLUCIA, INC.

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

SOLUCIA, INC.

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

61

SOLUCIA, INC.

Underwriting Decision-making

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

62

Underwriting Decision-making

SOLUCIA, INC.

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

Underwriting Decision-making

SOLUCIA, INC.

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

Underwriting Decision-making

SOLUCIA, INC.

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

Underwriting Decision-making

SOLUCIA, INC.

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

Underwriting Decision-making

SOLUCIA, INC.

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

Example 3: evaluating a high-risk model

SOLUCIA, INC.

Background

SOLUCIA, INC.

- 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.

Analysis

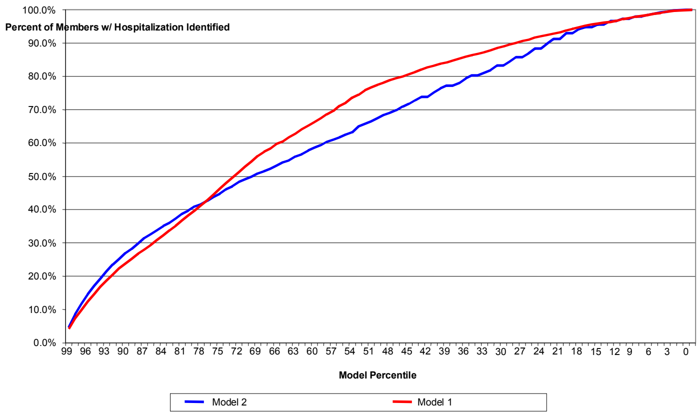
SOLUCIA, INC.

- 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.

Analysis

SOLUCIA, INC.

Lift Chart – Comparison between Two models



71

Analysis

SOLUCIA, INC.

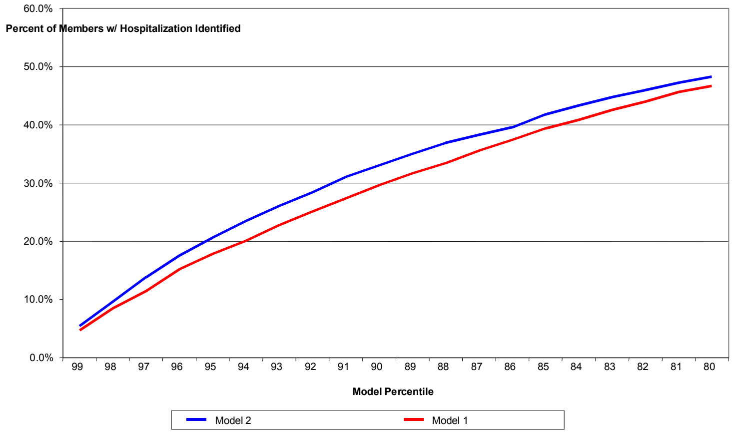
- Looked at over a narrower range, however, the results appear different.

72

Background

SOLUCIA, INC.

Lift Chart – Comparison between Two models



73

Analysis

SOLUCIA, INC.

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%	

74

Example 4: a wellness model

SOLUCIA, INC.

Solucia Wellness Model

SOLUCIA, INC.

- 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?

SOLUCIA, INC.

79

Selected references

SOLUCIA, INC.

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

80

Selected references (contd.)

SOLUCIA, INC.

- 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?

SOLUCIA, INC.

iduncan@soluciaconsulting.com

Solucia Consulting.
220 Farmington Avenue, Suite 4
Farmington, CT 06032

860-676-8808

www.soluciaconsulting.com