SOLUCIA, INC.

Introduction to Predictive Modeling

December 13, 2007

Introduction / Objective

SOLUCIA, INC.

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

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

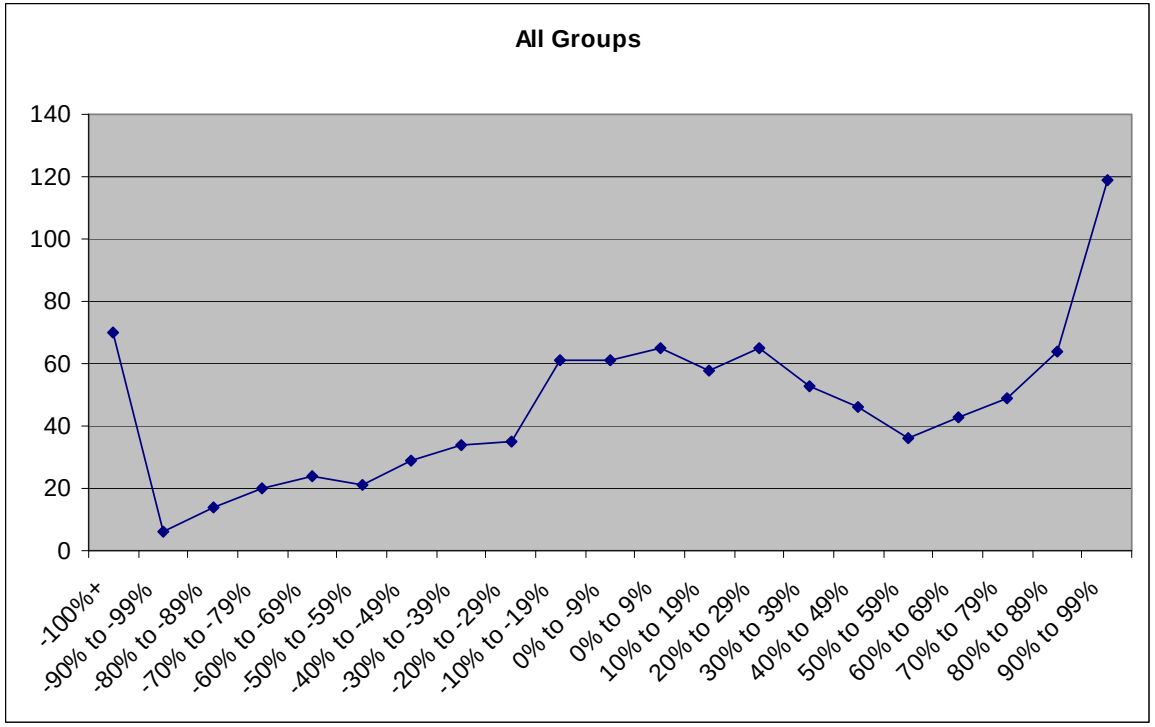
Background – Case 1

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

SOLUCIA, INC.

How well does the model perform?

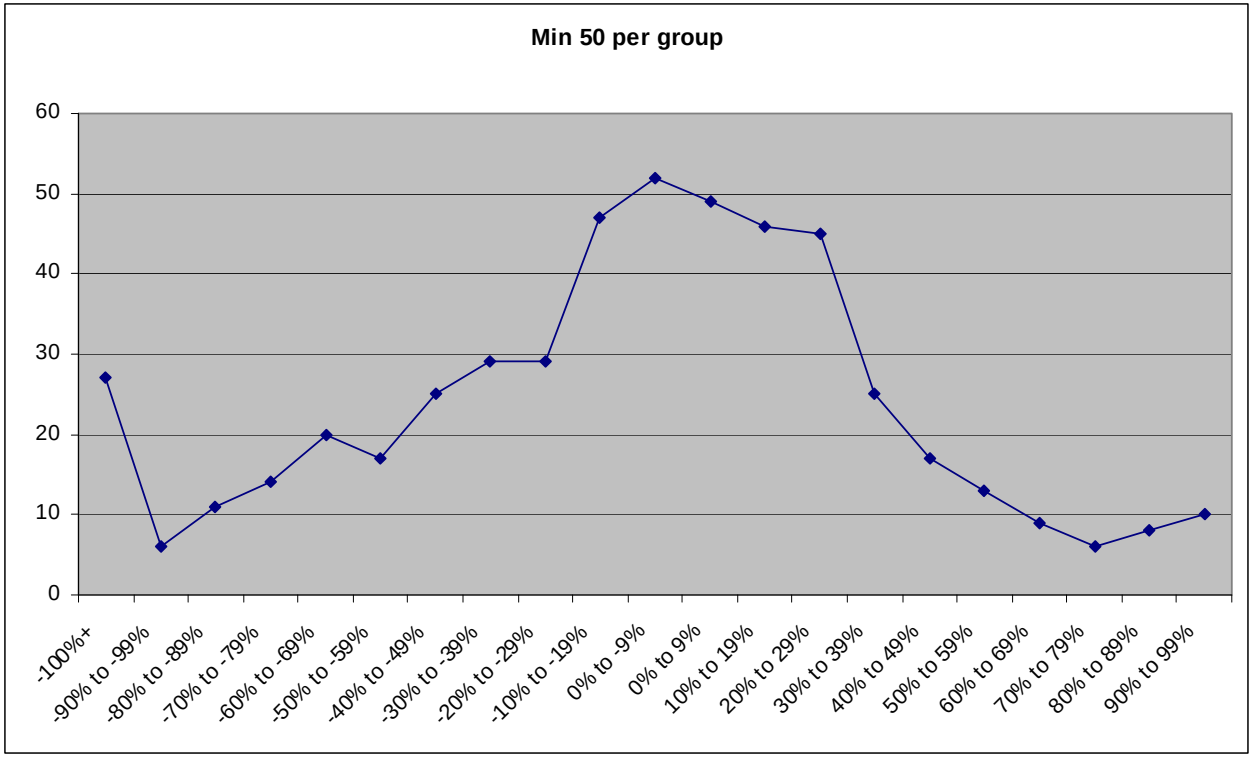
SOLUCIA, INC.



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.

How well does the model perform?

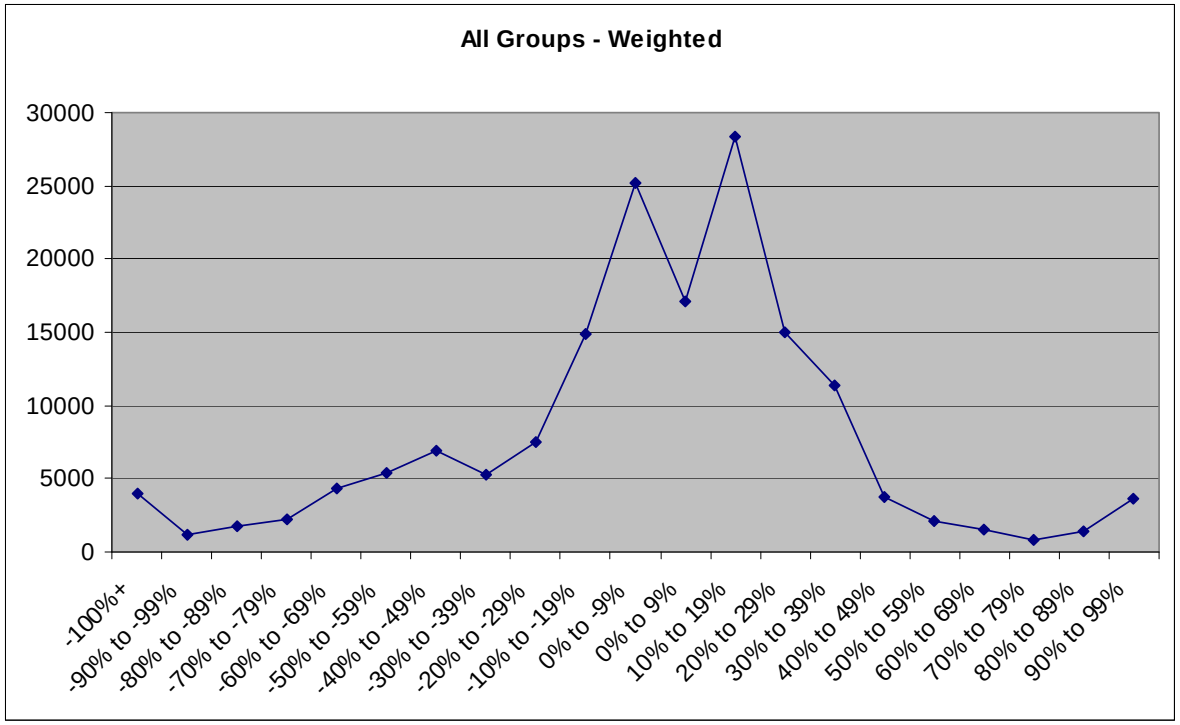
SOLUCIA, INC.



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

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.

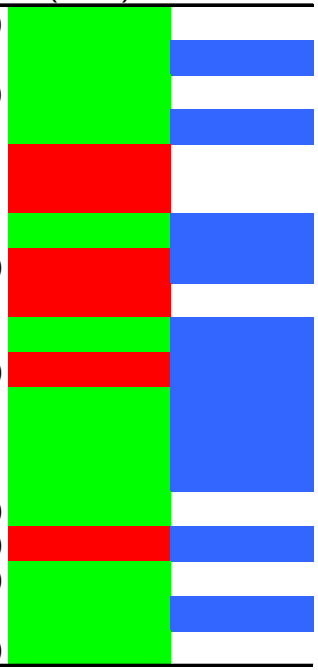
Node	PREDICTED Average Profit	PREDICTED Number in Node	PREDICTED Number in Node (Adjusted)	ACTUAL Number in node	ACTUAL Average Profit	Directionally Correct (+ or -)
1	(3.03)	70	173	170	(0.60)	
2	0.19	860	2,122	2,430	0.07	
3	(0.20)	2,080	5,131	6,090	(0.06)	
4	0.09	910	2,245	2,580	0.10	
5	(0.40)	680	1,678	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

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 -)	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		



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

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

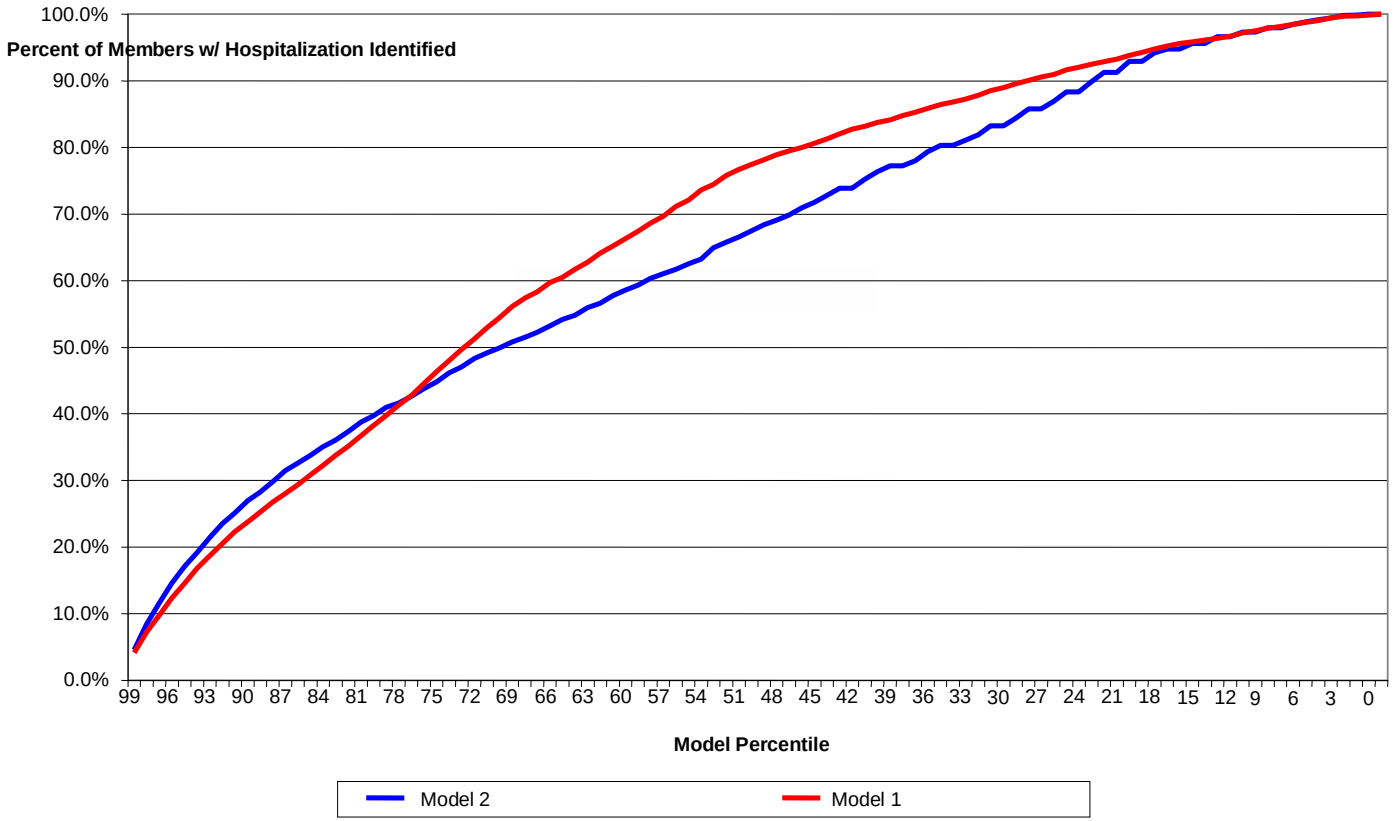
SOLUCIA, INC.

- 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



Analysis

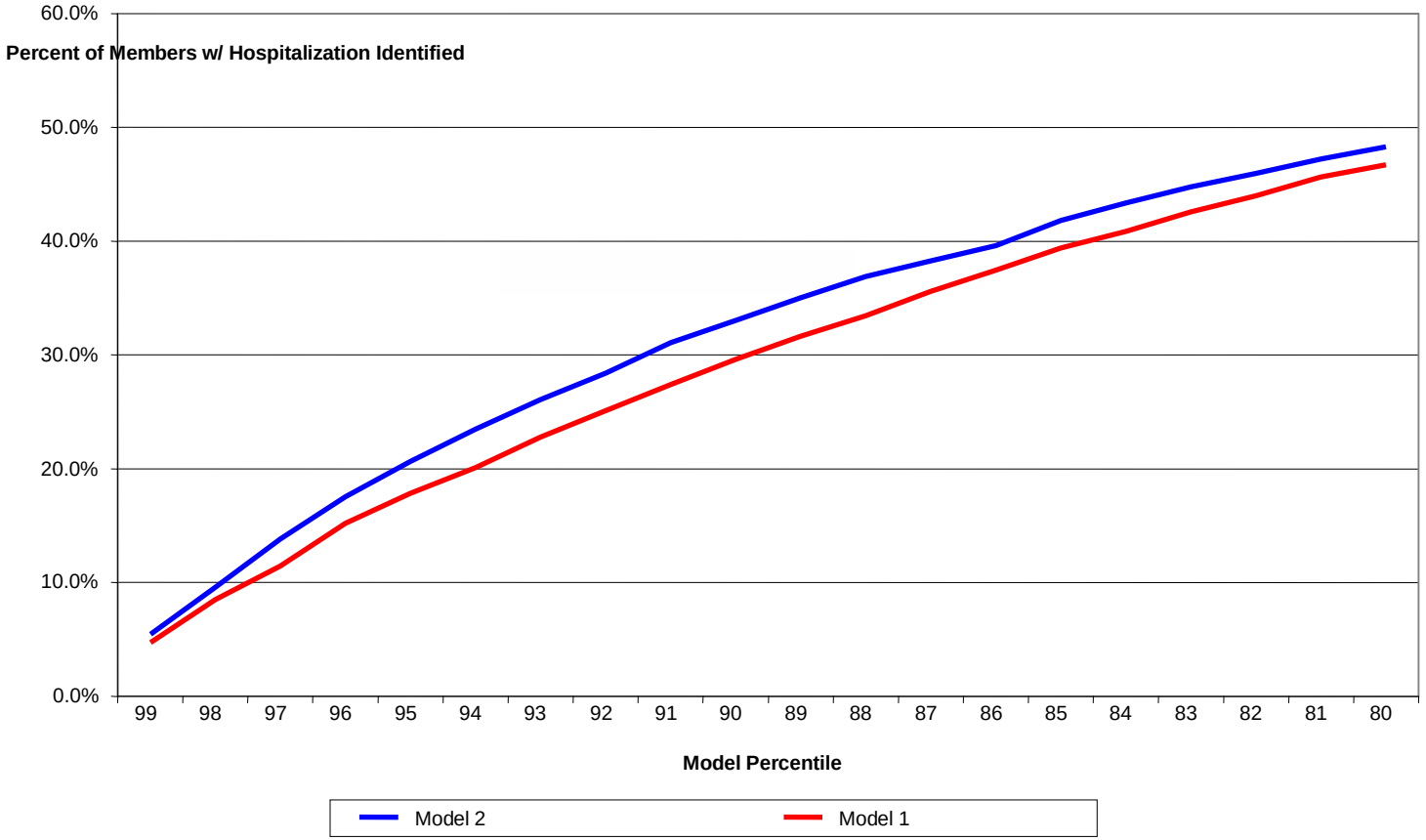
SOLUCIA, INC.

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

Background

SOLUCIA, INC.

Lift Chart – Comparison between Two models



Analysis

SOLUCIA, INC.

- Another way of looking at this….

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%	

Analysis

SOLUCIA, INC.

- And another ….

Analysis

Member Count Matrix

Model 1 Score Decile	Model 2 Score Decile 0	1	2	3	4	5	6	7	8	9	Grand Total
0	957	335	187	101	58	20	21	7	2	0	1688
1	383	484	319	224	134	77	38	24	6	0	1689
2	149	349	352	294	249	166	85	40	5	0	1689
3	73	196	304	293	326	219	156	93	27	2	1689
4	44	106	200	274	301	306	251	152	51	3	1688
5	24	71	170	196	276	295	303	238	106	12	1691
6	30	36	82	139	193	284	303	328	256	36	1687
7	21	31	57	84	130	180	291	388	393	114	1689
8	18	11	46	48	62	93	164	291	547	410	1690
9	9	5	12	14	31	41	61	96	306	1113	1688
Grand Total	1708	1624	1729	1667	1760	1681	1673	1657	1699	1690	16888

Analysis

Hospitalization Count Matrix

Model 1 Score Decile	Model 2 Score Decile 0	1	2	3	4	5	6	7	8	9	Grand Total
0	16	8	9	6	1	2	0	0	0	0	42
1	7	19	13	14	6	2	2	1	1	0	65
2	4	9	20	15	15	7	4	3	1	0	78
3	0	11	7	14	34	14	13	8	1	0	102
4	4	8	14	19	24	34	24	13	7	2	149
5	2	4	12	18	19	35	35	29	14	2	170
6	0	5	5	13	24	36	42	48	44	10	227
7	2	3	9	13	25	18	43	67	69	37	286
8	1	1	3	5	10	15	21	54	118	131	359
9	1	0	0	1	7	5	7	24	66	512	623
Grand Total	37	68	92	118	165	168	191	247	321	694	2101

Analysis

Hospitalization Rate Matrix

Model 1 Score Decile	Model 2 Score Decile										
	0	1	2	3	4	5	6	7	8	9	Grand Total
0	2%	2%	5%	6%	2%	10%	0%	0%	0%		2%
1	2%	4%	4%	6%	4%	3%	5%	4%	17%		4%
2	3%	3%	6%	5%	6%	4%	5%	8%	20%		5%
3	0%	6%	2%	5%	10%	6%	8%	9%	4%	0%	6%
4	9%	8%	7%	7%	8%	11%	10%	9%	14%	67%	9%
5	8%	6%	7%	9%	7%	12%	12%	12%	13%	17%	10%
6	0%	14%	6%	9%	12%	13%	14%	15%	17%	28%	13%
7	10%	10%	16%	15%	19%	10%	15%	17%	18%	32%	17%
8	6%	9%	7%	10%	16%	16%	13%	19%	22%	32%	21%
9	11%	0%	0%	7%	23%	12%	11%	25%	22%	46%	37%
Grand Total	2%	4%	5%	7%	9%	10%	11%	15%	19%	41%	12%

Evaluating the model

SOLUCIA, INC.

- The overlap/non-overlap between the LOH and All-encounter DxCG models. In the first table, 1,113 members are identified in the highest decile by both ICM and regular DxCG. The row numbers show the distribution of the balance of members, identified by the DxCG model, but assigned elsewhere by LOH. The column numbers show the DCG assignment of members with different LOH scores. Thus, for example, 410 members are assigned to decile 9 by the LOH model but to decile 8 by the DCG model. Conversely, 306 members assigned to decile 9 by the DCG model are assigned to 8 by the LOH model.
- Table 2 shows the actual admissions experienced within the different cells.
- Table 3 shows the hospitalization rates. Thus, for example, most of the DCG members in low deciles who are classified as decile 9 by the LOH model have a high admission rate (between 17% and 67%, depending on the DCG decile – on average 32% compared with 46% for the overlapping members in the top decile). On the other hand, the members in decile 9 by DCG score assigned to a lower decile by the LOH model have average admission rates between 7% and 25%, or an average of 19%).

Discussion?

SOLUCIA, INC.

Further Questions?

SOLUCIA, INC.

iduncan@soluciaconsulting.com

Solucia Inc.
220 Farmington Avenue
Farmington, CT 06032

860-676-8808

www.soluciaconsulting.com